

## POLISH MODELS AND SOFIC ENTROPY

BEN HAYES

**ABSTRACT.** We deduce properties of the Koopman representation of a positive entropy probability measure-preserving action of a countable, discrete, sofic group. Our main result may be regarded as a “representation-theoretic” version of Sinai’s Factor Theorem. We show that probability measure-preserving actions with completely positive entropy of an infinite sofic group must be mixing and, if the group is nonamenable, have spectral gap. This implies that if  $\Gamma$  is a nonamenable group and  $\Gamma \curvearrowright (X, \mu)$  is a probability measure-preserving action which is not strongly ergodic, then no action orbit equivalent to  $\Gamma \curvearrowright (X, \mu)$  has completely positive entropy. Crucial to these results is a formula for entropy in the presence of a Polish, but a priori noncompact, model.

## CONTENTS

1. Introduction	1
2. Preliminaries	6
2.1. Notational Remarks	6
2.2. Preliminaries on Sofic Groups	7
3. Definition of Entropy in the Presence of a Polish Model	9
4. Spectral Consequences of Positive Entropy	19
4.1. Representation Theoretic Preliminaries	19
4.2. Proofs of the Main Applications	21
References	25

## 1. INTRODUCTION

This paper is concerned with the structural consequences of positive entropy of probability measure-preserving actions of groups. Measure-theoretic entropy for actions of  $\mathbb{Z}$  is classical and goes back to the work of Kolmogorov and Sinai. Roughly speaking, it measures the randomness of the action of  $\mathbb{Z}$ . It was realized by Kieffer in [27] that one could replace  $\mathbb{Z}$  with the weaker condition of amenability of the group. Amenability requires a sequence of non-empty finite subsets of the group over which one can average in an approximately translation invariant way. Abelian groups, nilpotent groups and solvable groups are all amenable, whereas the free group on  $r$  letters is not if  $r \geq 1$ . Entropy for amenable groups is well established as a useful quantity in ergodic theory: it can be computed in many interesting cases (although it is not easy in general), when it is positive it reveals interesting structure on the action, and it has some useful general and intuitive properties.

The most fundamental examples in entropy theory are Bernoulli shifts. If  $\Gamma$  is a countable, discrete group and  $(\Omega, \omega)$  is a standard probability space the *Bernoulli shift* over  $(\Omega, \omega)$  is the probability measure-preserving action  $\Gamma \curvearrowright (\Omega, \omega)^\Gamma$  defined by

$$(gx)(h) = x(g^{-1}h), \quad x \in \Omega^\Gamma, g, h \in \Gamma.$$

Bernoulli shifts over amenable groups are completely classified by their entropy. If  $\Gamma$  is an infinite, amenable group and  $\Gamma \curvearrowright (X, \mu)$  is a free, ergodic, probability measure-preserving action with positive entropy, then  $\Gamma \curvearrowright (X, \mu)$  factors onto a Bernoulli shift. In fact, in this situation,  $\Gamma \curvearrowright (X, \mu)$  factors onto any Bernoulli shift of entropy at most that of  $\Gamma \curvearrowright (X, \mu)$ . This is known as Sinai’s factor theorem and it was proved for  $\Gamma = \mathbb{Z}$  by Sinai in [42] and for general amenable groups by Ornstein-Weiss in [35]. Sinai’s factor theorem

---

*Date:* January 5, 2016.

*2010 Mathematics Subject Classification.* 37A35, 47A67, 37A55, 46L89,

*Key words and phrases.* Sofic entropy, Sinai’s factor theorem, noncommutative harmonic analysis.

is a fundamental result in entropy theory, as it shows that Bernoulli factors “capture” all the entropy of a probability measure-preserving action of an amenable group and in some sense shows that entropy is simply a measure of the amount of “Bernoulli-like” behavior an action has.

In groundbreaking work, L. Bowen defined entropy for probability measure-preserving actions of sofic groups assuming the existence of a finite generating partition (see [5]). This assumption was then removed by Kerr-Li who also defined entropy for actions of sofic groups on compact, metrizable spaces (see [25]). We refer the reader to see Section 2.2 for the precise definition. Sofic groups form a vastly larger class of groups than amenable groups, it is known that amenable groups, free groups, residually finite groups and linear groups are all sofic and soficity is closed under free products with amalgamation over amenable subgroups (see [16],[15],[17],[36],[38]). Thus sofic entropy is a considerable extension of entropy for actions of an amenable group defined by Kieffer. Roughly, a group is sofic if it has “almost actions” on finite sets which are “almost free.” A sequence of such “almost actions” is called a sofic approximation. Entropy for a probability measure-preserving action  $\Gamma \curvearrowright (X, \mu)$  of a sofic group is then defined as the exponential growth rate of the number of finitary models there are of the action which are compatible with the fixed sofic approximation. Using sofic entropy, Bowen showed that if two Bernoulli shifts over a sofic group are isomorphic, then their base entropies are the same. Kerr-Li reproved this result with a more direct proof when the base space has infinite entropy in [24].

Since the subject is fairly young there are relatively few known structural consequences of positive *measure* entropy for actions of *arbitrary* sofic groups. For example, Sinai’s factor theorem is not known for sofic groups. Previous consequences of entropy for actions of sofic groups have either been for topological actions or for specific groups. In [26], Kerr-Li prove that actions with positive *topological* entropy must exhibit some chaotic behavior (for example, they must be Li-Yorke chaotic). When the acting group is a free group one can consider another form of measure entropy, defined by Bowen in [6], called *f*-invariant entropy. The *f*-invariant entropy is roughly a “randomized” version of sofic entropy. Interesting consequences have been given by Seward in [40] for the case of *f*-invariant entropy, but those only apply when the group is a free group. After the appearance of our preprint, Meyerovitch showed in [33] that positive sofic entropy implies that almost every stabilizer of the action is finite.

In this paper, we will deduce structural consequences of positive measure entropy for actions of arbitrary sofic groups. Our applications are to *spectral* properties of such actions. To the best of our knowledge, aside from the results in our paper, the results of Meyerovitch are the only ones which deduce properties of an action of a *general* sofic group assuming only that the action has positive entropy. Recall that if  $\Gamma \curvearrowright (X, \mu)$  is a probability measure-preserving action of a countable discrete group  $\Gamma$ , we have an induced unitary representation  $\rho_{\Gamma \curvearrowright (X, \mu)} : \Gamma \rightarrow \mathcal{U}(L^2(X, \mu))$  given by

$$(\rho_{\Gamma \curvearrowright (X, \mu)}(g)\xi)(x) = \xi(g^{-1}x).$$

We use  $\rho_{\Gamma \curvearrowright (X, \mu)}^0$  for the restriction of  $\rho_{\Gamma \curvearrowright (X, \mu)}$  to  $L^2(X, \mu) \ominus \mathbb{C}1$ . The representation  $\rho_{\Gamma \curvearrowright (X, \mu)}^0$  is called the *Koopman representation* of  $\Gamma \curvearrowright (X, \mu)$ . Properties of a probability-measure preserving action which only depend upon the Koopman representation are called *spectral* properties. The Koopman representation has played a significant role in ergodic theory since the early days of the subject, being the means to deduce von Neumann’s Mean Ergodic Theorem which the Ergodic Theorem relies upon as a first step. Additionally, many other fundamental properties such as compactness, weak mixing, mixing and ergodicity are all spectral properties.

Our results show that one canonical representation of a group plays a special role in entropy theory. Recall that the left regular representation of a group  $\lambda : \Gamma \rightarrow \mathcal{U}(\ell^2(\Gamma))$  is defined by  $(\lambda(g)f)(x) = f(g^{-1}x)$ . We will sometimes use  $\lambda_\Gamma$  to specify the group. We will also use  $\lambda_{\Gamma, \mathbb{R}}$  (or  $\lambda_{\mathbb{R}}$  if the group is clear) for the orthogonal representation which is the restriction of  $\lambda_\Gamma$  to  $\ell^2(\Gamma, \mathbb{R})$ .

Let  $\rho_j : \Gamma \rightarrow \mathcal{U}(\mathcal{H}_j)$  be two unitary representations of a countable discrete group  $\Gamma$ . We say that  $\rho_1, \rho_2$  are *singular* and write  $\rho_1 \perp \rho_2$  if no nonzero subrepresentation of  $\rho_1$  embeds into  $\rho_2$ . By subrepresentation, we mean a restriction of  $\rho_j, j = 1, 2$  to a closed,  $\Gamma$ -invariant, linear subspace. As is customary, we will often call a closed,  $\Gamma$ -invariant, linear subspace a subrepresentation as well. The terminology of singularity comes from the case  $\Gamma = \mathbb{Z}$ . If  $\mu$  is a Borel measure on  $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ , then we have a natural unitary representation

$$\rho_\mu : \mathbb{Z} \rightarrow \mathcal{U}(L^2(\mathbb{T}, \mu))$$

given by

$$(\rho_\mu(n)\xi)(\theta) = e^{2\pi i n \theta} \xi(\theta).$$

It is easy to check that  $\rho_\mu \perp \rho_\nu$  if and only if  $\mu \perp \nu$ . Similar analysis can be done for any abelian group (replacing  $\mathbb{T}$  by the Pontryagin dual). Thus, singularity of representations is a natural generalization to noncommutative groups of singularity of measures.

If  $\Gamma \curvearrowright (X, \mu)$  is a probability measure-preserving action of a countable, discrete sofic group and  $\Sigma$  is a sofic approximation of  $\Gamma$  (see Definition 2.1 for the precise definition of a sofic approximation) we use  $h_{\Sigma, \mu}(X, \Gamma)$  for the entropy of  $\Gamma \curvearrowright (X, \mu)$  with respect to  $\Sigma$  as defined by Bowen, Kerr-Li.

**Theorem 1.1.** *Let  $\Gamma$  be a countably infinite discrete sofic group with sofic approximation  $\Sigma$ . Let  $\Gamma \curvearrowright (X, \mathcal{M}, \mu)$  be a measure-preserving action where  $(X, \mathcal{M}, \mu)$  is a standard probability space. Suppose that  $\mathcal{H} \subseteq L^2(X, \mu)$  is a subrepresentation such that  $\mathcal{M}$  is generated (up to sets of measure zero) by*

$$\{f^{-1}(A) : f \in \mathcal{H}, A \subseteq \mathbb{C} \text{ is Borel}\}.$$

*If  $\rho_{\Gamma \curvearrowright (X, \mu)}^0|_{\mathcal{H}}$  is singular with respect to the left regular representation, then*

$$h_{\Sigma, \mu}(X, \Gamma) \leq 0.$$

It is well-known that if  $(\Omega, \omega)$  is a standard probability space, then for the Bernoulli action  $\Gamma \curvearrowright (\Omega, \omega)^\Gamma$  we have

$$(1) \quad \rho_{\Gamma \curvearrowright (\Omega, \omega)^\Gamma}^0 \cong \lambda_\Gamma^{\oplus \infty}.$$

We mentioned before that Sinai's factor theorem is not known for sofic groups. Note that if  $\Gamma \curvearrowright (X, \mu)$  factors onto a Bernoulli shift, then by (1) we have that  $\lambda_\Gamma^{\oplus \infty}$  embeds into  $\rho_{\Gamma \curvearrowright (X, \mu)}^0$ . In this manner, Theorem 1.1 may be regarded as a weak version of Sinai's Factor Theorem for sofic groups. It shows that at the representation-theoretic level, an action of a sofic group with positive entropy must contain a subrepresentation of the Koopman representation of a Bernoulli shift. We can thus think of this theorem as a "representation-theoretic version" of Sinai's Factor Theorem. It is also the first result which indicates that positive entropy actions of sofic groups must behave in a manner "similar" to Bernoulli shifts.

We can say even more than Theorem 1.1 if we assume a stronger version of positive entropy. Recall that an action  $\Gamma \curvearrowright (X, \mu)$  has *completely positive entropy* with respect to a sofic approximation  $\Sigma$  if whenever  $\Gamma \curvearrowright (Y, \nu)$  is a factor of  $\Gamma \curvearrowright (X, \mu)$  and  $(Y, \nu)$  is not a one-atom space then  $h_{\Sigma, \nu}(Y, \Gamma) > 0$ . The following is easy from Theorem 1.1.

**Corollary 1.2.** *Let  $\Gamma$  be a countable discrete sofic group with sofic approximation  $\Sigma$ . Suppose that  $\Gamma \curvearrowright (X, \mu)$  has completely positive entropy with respect to  $\Sigma$ . Then the Koopman representation of  $\Gamma \curvearrowright (X, \mu)$  is embeddable into the infinite direct sum of the left regular representation.*

Corollary 1.2 was proved for  $\Gamma$  amenable by Dooley and Golodets in [14]. Here we should mention that Dooley-Golodets actually prove that if  $\Gamma \curvearrowright (X, \mu)$  has completely positive entropy and  $\Gamma$  is amenable, then the Koopman representation of  $\Gamma \curvearrowright (X, \mu)$  is isomorphic to an infinite direct sum of the left regular representation. This is an easy consequence of Corollary 1.2 and Sinai's factor theorem. Since Sinai's factor theorem is not known for sofic groups, we do not know if  $\rho_{\Gamma \curvearrowright (X, \mu)}^0 \cong \lambda^{\oplus \infty}$  (instead of just  $\rho_{\Gamma \curvearrowright (X, \mu)}^0$  embeds into  $\lambda^{\oplus \infty}$ ) when  $\Gamma$  is sofic.

From Corollary 1.2, we automatically deduce other important structural properties of completely positive entropy actions of a sofic group, specifically mixing and spectral gap. A probability measure-preserving action  $\Gamma \curvearrowright (X, \mu)$  is said to:

- (i): be mixing if  $\lim_{g \rightarrow \infty} \mu(gA \cap B) = \mu(A)\mu(B)$  for all measurable  $A, B \subseteq X$ ,
- (ii): be strongly ergodic if for every sequence  $A_n$  of measurable subsets of  $X$  with  $\mu(gA_n \Delta A_n) \rightarrow 0$  for all  $g \in \Gamma$  we have that  $\mu(A_n)(1 - \mu(A_n)) \rightarrow 0$ ,
- (iii): have spectral gap if for every sequence  $\xi_n \in L^2(X, \mu)$  with  $\|g\xi_n - \xi_n\|_2 \rightarrow 0$  for all  $g \in \Gamma$ , it is true that  $\|\xi_n\|_2 \rightarrow 0$ .

It is easy to see that spectral gap implies strong ergodicity.

**Corollary 1.3.** *Let  $\Gamma$  be a countable discrete sofic group with sofic approximation  $\Sigma$ . Suppose that  $\Gamma \curvearrowright (X, \mu)$  has completely positive entropy with respect to  $\Sigma$ .*

*(i): If  $\Gamma$  is infinite, then  $\Gamma \curvearrowright (X, \mu)$  is mixing.*

*(ii): If  $\Lambda$  is any nonamenable subgroup of  $\Gamma$ , then  $\Lambda \curvearrowright (X, \mu)$  is strongly ergodic (in fact it has spectral gap).*

There is another approach to nonamenable entropy, called Rokhlin entropy, due to Seward in [41]. It has the advantage of being easy to define and being defined for all groups, but the disadvantage of being extremely difficult to compute. Rokhlin entropy is an upper bound for sofic entropy, but there are no known cases where one can prove an action has positive Rokhlin entropy without using that it has positive sofic entropy. We mention that completely positive Rokhlin entropy can be defined in a similar manner. After the appearance of our preprint, Alpeev in [1] proved that actions with completely positive Rokhlin entropy are weakly mixing. His approach is completely elementary. It is easy to see that completely positive sofic entropy implies completely positive Rokhlin entropy. Whether or not actions of a nonamenable group with completely positive Rokhlin entropy are mixing, strongly ergodic, or have spectral gap all appear to be open. It seems to be very difficult to deduce properties of the Koopman representation of an action from the assumption that the action has positive Rokhlin entropy.

Part (ii) of the above Corollary is rather special to nonamenable groups. Recall that two probability measure-preserving actions  $\Gamma \curvearrowright (X, \mu), \Lambda \curvearrowright (Y, \nu)$  of countable discrete groups  $\Gamma, \Lambda$  are said to be *orbit equivalent* if there is a measure space isomorphism  $\theta: (X, \mu) \rightarrow (Y, \nu)$  so that  $\theta$  takes the  $\Gamma$ -orbits to  $\Lambda$ -orbits, i.e.  $\theta(\Gamma x) = \Lambda\theta(x)$  for almost every  $x \in X$ . Orbit equivalence theory is an area of much current interest relating to operator algebras, ergodic theory and group theory. Strong ergodicity is an invariant of the orbit equivalence class of the action. Our corollary shows that if  $\Gamma \curvearrowright (X, \mu)$  is a probability measure-preserving action of a nonamenable group and if the action is not strongly ergodic, then no action orbit equivalent to  $\Gamma \curvearrowright (X, \mu)$  has completely positive entropy. Our result indicates that entropy for actions of nonamenable groups may have nontrivial consequences for orbit equivalence theory. It is a celebrated, and deep, fact that all ergodic actions of an amenable group are orbit equivalent. This is due to Connes-Feldman-Weiss and Ornstein-Weiss (see [12],[34]). Thus entropy for actions of amenable groups cannot have *any* consequences for orbit equivalence theory.

Spectral gap and strong ergodicity are important properties with many applications. Spectral gap has connections to expander graphs (see [3],[2],[4], [31]), orbit equivalence rigidity (see [21],[37]) and number theory (see [39]). Spectral gap is also related to the Banach-Ruziewicz problem which asks if Lebesgue measure is the unique, finitely-additive, rotation-invariant, probability measure on the sphere defined on all Lebesgue measurable sets (solved independently by Margulis [32] and Sullivan [43]). It is well known that *no* action of an amenable group is strongly ergodic. The consequences entropy has for strong ergodicity, spectral gap and orbit equivalence indicate that entropy for nonamenable groups may be used to deduce phenomena not present in the realm of ergodic theory of amenable groups. This reveals the importance of generalizing entropy to actions of nonamenable groups.

We briefly outline the key differences in our approach and the approach of Dooley-Golodets to prove Corollary 1.2. Dooley-Golodets first prove Theorem 1.1 when  $\Gamma = \mathbb{Z}$ . Fourier analysis reduces Theorem 1.1 for  $\Gamma = \mathbb{Z}$  to the fact that if  $\mu, \nu$  are mutually singular, Borel, probability measures on the circle, then there is a continuous function  $f$  on the circle with  $0 \leq f \leq 1$  so that  $f$  is “close” to the constant function 1 in  $L^2(\mu)$ , but is “close” to zero in  $L^2(\nu)$ . This fact is a simple exercise in measure theory. Dooley-Golodets then deduce Corollary 1.2 for  $\Gamma$  amenable from the case  $\Gamma = \mathbb{Z}$  by using that all amenable groups are orbit equivalent to the integers.

Our approach is to prove Theorem 1.1 for a general sofic group, by simply replacing the harmonic analysis for  $\Gamma = \mathbb{Z}$  with noncommutative harmonic analysis for a general group. The assumption about singularity of measures is then replaced by singularity of representations of groups. The representation theory of a group is captured by its universal  $C^*$ -algebra and so it is natural to replace the algebra of continuous functions on  $\mathbb{T}$  with the  $C^*$ -algebra of the group. One can then characterize singularity of representations of an arbitrary group in a manner similar to the preceding paragraph. In short, we abstract the harmonic analysis for the case  $\Gamma = \mathbb{Z}$  to noncommutative harmonic analysis for a nonabelian group. This approach uses essentially no structure of the group and removes the orbit equivalence techniques in the approach of [14], which are only valid in the case of amenable groups.

Although this approach using noncommutative harmonic analysis may make our methods seem abstract and esoteric, the proof of Theorem 1.1 is essentially self-contained. After the input of the aforementioned noncommutative harmonic analysis techniques, as well as basic facts about Borel measures on Polish spaces, the rest of the proof of Theorem 1.1 is elementary. It only relies on basic consequences of the finite-dimensional spectral theorem and simple volume counting estimates. Additionally, the noncommutative harmonic analysis techniques used lie at the basics of  $C^*$ -algebra theory. On the other hand, the fact that every action of amenable groups is orbit equivalent to the integers is fairly deep. So in the amenable case we have discovered a more direct and elementary proof of Corollary 1.2.

Intuitively, entropy should be some measure of randomness of the system. Theorem 1.1 shows indeed that positive entropy actions must exhibit some randomness properties. For example, we can view the left regular representation as the representation which exhibits the perfect amount of mixing. We can also use Theorem 1.1 to show that highly structured actions, like compact actions, must have nonpositive sofic entropy.

**Corollary 1.4.** *Let  $\Gamma$  be a countable discrete sofic group with sofic approximation  $\Sigma$ . Suppose that  $\alpha: \Gamma \rightarrow \text{Aut}(X, \mu)$  has image contained in a compact (for the weak topology) subgroup of  $\text{Aut}(X, \mu)$ . Then for the action given by  $\alpha$ ,*

$$h_{\Sigma, \mu}(X, \Gamma) \leq 0.$$

In a previous version of this article, we used our techniques to prove that distal measure-theoretic action have entropy zero. After the appearance of this version, Alpeev proved in [1] that measure distal actions of an arbitrary group have Rohklin entropy zero. Then, Burton proved in [10] that distal actions have naive entropy zero if the group contains a copy of  $\mathbb{Z}$ . Naive entropy zero implies Rohklin entropy zero which in turn implies sofic entropy zero. These proofs are both elementary, while ours is arguably not. We have elected to remove this section from the current version of this article, as we are actually going to prove more general results in [19] which will give structural results of a probability measure-preserving action of a sofic group relative to its Pinsker factor. The results in [1] imply that Rohklin entropy decreases for compact extensions (and our results in [19] prove the same result for sofic entropy) and it appears that it is unknown whether or not naive entropy decreases under compact extensions.

An important new tool we use to prove all of these theorems is Polish models. If  $\Gamma \curvearrowright (X, \mu)$  is a probability measure-preserving action of a countable discrete group, a *topological model* for the action is an action  $\Gamma \curvearrowright (Y, \nu)$  isomorphic to  $\Gamma \curvearrowright (X, \mu)$  where  $Y$  is a separable, metrizable topological space,  $\nu$  is a  $\Gamma$ -invariant Borel probability measure on  $Y$  and the action is by homeomorphisms. Roughly one can think of this as giving a topology on  $X$  so that the action is by homeomorphisms. For actions on standard probability spaces compact models always exist. Moreover, Kerr-Li in [25] show that one can compute entropy in the presence of a compact model in a manner which uses the topology. Many of the computations for sofic entropy have used the compact model formalism (see [28], [25], [18]). To prove the above theorems, we give a definition of sofic entropy in the presence of a Polish model (that is where  $Y$  is merely assumed to be a Polish space, i.e. a completely metrizable separable topological space). We remark here that R. Bowen defined (see [8]) topological entropy for uniformly continuous automorphisms of a metric space, proving that it was invariant under uniformly continuous conjugacies. Our approach is slightly different here, we do not require the homeomorphisms to be uniformly continuous and R. Bowen did not consider measure-preserving actions.

Since compact models always exist, we should mention why we decided to consider the case of a Polish model. For this, let us mention a natural way of obtaining Polish models. Given a probability measure-preserving action  $\Gamma \curvearrowright (X, \mathcal{M}, \mu)$ , we say that a family  $\mathcal{F}$  of measurable functions  $X \rightarrow \mathbb{C}$  is *generating* if  $\mathcal{M}$  is the smallest complete,  $\Gamma$ -invariant sigma-algebra of sets in  $X$  which makes all the elements of  $\mathcal{F}$  measurable. Associated to a family of generators, one can canonically produce a topological model in the following manner. Define

$$\Phi: X \rightarrow \mathbb{C}^{\mathcal{F} \times \Gamma}$$

by

$$\Phi(x)(f, g) = f(g^{-1}x)$$

and let  $\Gamma \curvearrowright \mathbb{C}^{\mathcal{F} \times \Gamma}$  be the shift action given by

$$(gx)(f, h) = x(f, g^{-1}h).$$

Setting  $\nu = \Phi_*\mu$ , we have that  $\Gamma \curvearrowright (\mathbb{C}^{\mathcal{F} \times \Gamma}, \nu)$  is a topological model of our action. We can thus think of topological models as a “presentation-theoretic” approach to ergodic theory analogous to presentation theory of groups. If  $\mathcal{F} \subseteq L^\infty(X, \mu)$ , then the topological model one produces above is a compact model. In general the topological model is Polish. Thus Polish models are a canonical way of dealing with unbounded generators. This is relevant because all of our assumptions are about functions in  $L^2(X, \mu)$  and not  $L^\infty(X, \mu)$ . There are of course ways of turning a family of generators which are unbounded into family of generators which are bounded. One can employ cut-off functions, or compose with injective continuous maps  $\mathbb{R} \rightarrow [-1, 1]$ . We warn the reader that *all* of these attempts to reduce to a family of bounded generators destroy our representation-theoretic hypotheses and so we will really need to deal with unbounded generators and, as a consequence, Polish models.

We remark here that if we take the above topological model associated to family of generators in  $L^\infty(X, \mu)$  then we essentially recover the operator algebraic approach to sofic entropy by Kerr-Li in [25]. Thus the Polish model approach to sofic entropy may be regarded as a generalization of the operator algebra approach given by Kerr-Li in [25] to a family of unbounded generators. The crucial aspect of Polish spaces which allows us to equate our definition of entropy for a Polish model to that of Bowen and Kerr-Li is tightness of a single probability measure on a Polish space. Tightness roughly asserts that, up to a small error, the probability measure the probability measure is supported on a compact set. There are examples of separable metrizable spaces where not every Borel probability measure is tight (e.g. see the remarks after Theorem 23 of [44]) and this is the reason we need our spaces to be Polish. The assumption of a topological model being Polish is also natural since the canonical topological model associated to a family of generators is always Polish.

**Acknowledgments** Much of this work was done while I was still a PhD student at UCLA. I am very grateful to the kind hospitality and stimulating environment at UCLA. I would like to thank Lewis Bowen for suggesting the problem of computing sofic entropy of Gaussian actions in the program “Von Neumann algebras and ergodic theory of group actions” at Institut Henri Poincaré in 2011. The solution to this problem ultimately led to this work. I would like to thank Stephanie Lewkiewicz for many interesting discussion about Polish models.

## 2. PRELIMINARIES

**2.1. Notational Remarks.** We will use  $\rho$  for a representation. We will thus have to forego the usual practice in sofic entropy of using  $\rho$  for a metric, and will instead use  $\Delta$ .

If  $A, B$  are sets we use  $B^A$  for all functions  $f: A \rightarrow B$ . If  $A = \{1, \dots, n\}$  we will use  $B^n$  instead of  $B^{\{1, \dots, n\}}$ . If  $C$  is another set and  $\phi: B \rightarrow C$  is a function, we use  $\phi^n: B^n \rightarrow C^n$  for the map

$$\phi^n(f) = f \circ \phi.$$

If  $X$  is a Polish space, and  $f \in C_b(X)$  we use  $\|f\|$  for the uniform norm of  $f$ . We will use  $\|f\|_{C_b(X)}$  if the space  $X$  is not clear from the context.

Let  $(A, \Delta)$  be a pseudometric space. For subsets  $C, B$  of  $A$ , and  $\varepsilon > 0$  we say that  $C$  is  $\varepsilon$ -contained in  $B$  and write  $C \subseteq_\varepsilon B$  if for all  $c \in C$ , there is a  $b \in B$  so that  $\Delta(c, b) < \varepsilon$ . We say that  $S \subseteq A$  is  $\varepsilon$ -dense if  $A \subseteq_\varepsilon S$ . We use  $S_\varepsilon(A, \Delta)$  for the smallest cardinality of a  $\varepsilon$ -dense subset of  $A$ . If  $C \subseteq_\delta B$  are subsets of  $A$ , then

$$S_{2(\varepsilon+\delta)}(C, \Delta) \leq S_\varepsilon(B, \Delta).$$

If  $A$  is a finite set we will use  $u_A$  for the uniform probability measure on  $A$ . We will typically write  $u_n$  instead of  $u_{\{1, \dots, n\}}$ . For  $x \in \mathbb{C}^n$ ,  $\|x\|_2$  will denote the  $\ell^2$  norm with respect to  $u_n$  unless otherwise stated. There are certain times when we will have to use  $\|\cdot\|_{\ell^2(n)}$ ,  $\|\cdot\|_{\ell^2(n, u_n)}$  at the same time. In such an instance we will use notation which specifies which  $\ell^2$ -norm we are using. Additionally, we will use

$$\langle \cdot, \cdot \rangle_{\ell^2(n)},$$

$$\langle \cdot, \cdot \rangle_{\ell^2(n, u_n)}$$

for the inner products on  $\ell^2(n)$ ,  $\ell^2(n, u_n)$  when there is potential confusion. If it is not otherwise specified, then

$$\langle \cdot, \cdot \rangle$$

refers to the inner product with respect to  $\ell^2(n, u_n)$ .

We say that  $N \subseteq A$  is  $\varepsilon$ -separated if for every  $n_1 \neq n_2$  in  $N$  we have  $\Delta(n_1, n_2) > \varepsilon$ . We use  $N_\varepsilon(A, \Delta)$  for the smallest cardinality of a  $\varepsilon$ -separated subset of  $A$ . Note that

$$(2) \quad N_{2\varepsilon}(A, \Delta) \leq S_\varepsilon(A, \Delta) \leq N_\varepsilon(A, \Delta),$$

and that if  $A \subseteq B$ , then

$$N_\varepsilon(A, \Delta) \leq N_\varepsilon(B, \Delta).$$

**2.2. Preliminaries on Sofic Groups.** We use  $S_n$  for the symmetric group on  $n$  letters. If  $A$  is a set, we will use  $\text{Sym}(A)$  for the set of bijections of  $A$ .

**Definition 2.1.** Let  $\Gamma$  be a countable discrete group. A *sofic approximation* of  $\Gamma$  is a sequence  $\Sigma = (\sigma_i : \Gamma \rightarrow S_{d_i})$  of functions (not assumed to be homomorphisms) so that

$$\lim_{i \rightarrow \infty} u_{d_i}(\{1 \leq k \leq d_i : \sigma_i(g)\sigma_i(h)(k) = \sigma_i(gh)(k)\}) = 1, \text{ for all } g, h \in \Gamma$$

$$\lim_{i \rightarrow \infty} u_{d_i}(\{1 \leq k \leq d_i : \sigma_i(g)(k) \neq \sigma_i(h)(k)\}) = 1, \text{ for all } g \neq h \text{ in } \Gamma.$$

We will call  $\Gamma$  *sofic* if it has a sofic approximation.

It is known that all amenable groups and residually finite groups are sofic. Also, it is known that soficity is closed under free products with amalgamation over amenable subgroups (see [17],[36],[16],[15], [38]). Also graph products of sofic groups are sofic by [11]. Additionally, residually sofic groups and locally sofic groups are sofic. Thus by Malcev's Theorem we know all linear groups are sofic. Finally, if  $\Lambda$  is a subgroup of  $\Gamma$ , and  $\Lambda$  is sofic and  $\Gamma \curvearrowright \Gamma/\Lambda$  is amenable (in the sense of having a  $\Gamma$ -invariant mean) then  $\Gamma$  is sofic (this can be seen by a mild generalization of the argument in Theorem 1 of [17] using e.g. the observation after Definition 12.2.12 of [9]).

We will need to extend a sofic approximation to certain algebras associated to  $\Gamma$ . Let  $\mathbb{C}(\Gamma)$  be the ring of finite formal linear combinations of elements of  $\Gamma$  with addition defined naturally and multiplication defined by

$$\left( \sum_{g \in \Gamma} a_g g \right) \left( \sum_{h \in \Gamma} b_h h \right) = \sum_{g \in \Gamma} \left( \sum_{h \in \Gamma} a_h b_{h^{-1}g} \right) g.$$

We will also define a conjugate-linear involution on  $\mathbb{C}(\Gamma)$  by

$$\left( \sum_{g \in \Gamma} a_g g \right)^* = \sum_{g \in \Gamma} \overline{a_{g^{-1}}} g.$$

Given a sofic approximation  $\Sigma = (\sigma_i : \Gamma \rightarrow S_{d_i})$  and  $\alpha = \sum_{g \in \Gamma} \alpha_g g \in \mathbb{C}(\Gamma)$  we define  $\sigma_i(\alpha) \in M_{d_i}(\mathbb{C})$  by

$$\sigma_i(\alpha) = \sum_{g \in \Gamma} \alpha_g \sigma_i(g).$$

In order to talk about the asymptotic properties of this extended sofic approximation, we will need a more analytic object associated to  $\Gamma$ .

Let  $\lambda : \Gamma \rightarrow \mathcal{U}(\ell^2(\Gamma))$  be the left regular representation defined by  $(\lambda(g)\xi)(h) = \xi(g^{-1}h)$ . We will continue to use  $\lambda$  for the linear extension to  $\mathbb{C}(\Gamma) \rightarrow B(\ell^2(\Gamma))$ . The *group von Neumann algebra* of  $\Gamma$  is defined by

$$\overline{\lambda(\mathbb{C}(\Gamma))}^{\text{WOT}}$$

where WOT denotes the weak operator topology. We will use  $L(\Gamma)$  to denote the group von Neumann algebra. Define  $\tau : L(\Gamma) \rightarrow \mathbb{C}$  by

$$\tau(x) = \langle x\delta_e, \delta_e \rangle.$$

We leave it as an exercise to the reader to verify that  $\tau$  has the following properties.

- 1:  $\tau(1) = 1$ ,
- 2:  $\tau(x^*x) \geq 0$ , with equality if and only if  $x = 0$ ,
- 3:  $\tau(xy) = \tau(yx)$ , for all  $x, y \in M$ ,
- 4:  $\tau$  is weak operator topology continuous.

We call the third property the tracial property. We will typically view  $\mathbb{C}(\Gamma)$  as a subset of  $L(\Gamma)$ . In particular, we will use  $\tau$  as well for the functional on  $\mathbb{C}(\Gamma)$  which is just the restriction of  $\tau$  on  $L(\Gamma)$ .

In order to state our extension of a sofic approximation properly, we shall give a general definition. Recall that  $*$ -algebra is a complex algebra equipped with an involution  $*$  which is conjugate linear and antimultiplicative.

**Definition 2.2.** A *tracial  $*$ -algebra* is a pair  $(A, \tau)$  where  $A$  is a  $*$ -algebra equipped with a linear functional  $\tau: A \rightarrow \mathbb{C}$  so that

$$1: \tau(1) = 1,$$

$$2: \tau(x^*x) \geq 0, \text{ with equality if and only if } x = 0,$$

$$3: \tau(xy) = \tau(yx), \text{ for all } x, y \in A,$$

$$4: \text{ For all } a \in A, \text{ there is a } C_a > 0 \text{ so that for all } x \in A, |\tau(x^*a^*ax)| \leq C_a\tau(x^*x).$$

For  $a, b \in A$  we let  $\langle a, b \rangle = \tau(b^*a)$  and we let  $\|a\|_2 = \tau(a^*a)^{1/2}$ . We let  $L^2(A, \tau)$  be the Hilbert space completion of  $A$  in this inner product. By condition 4 of the definition, we have a representation  $\lambda: A \rightarrow B(L^2(A, \tau))$  defined densely by  $\lambda(a)x = ax$  for  $x \in A$ . We let  $\|a\|_\infty = \|\lambda(a)\|$ .

We make  $M_n(\mathbb{C})$  into a tracial  $*$ -algebra using  $\text{tr} = \frac{1}{n} \text{Tr}$  where  $\text{Tr}$  is the usual trace. In particular, we use  $\|A\|_2 = \text{tr}(A^*A)^{1/2}$  and  $\|A\|_\infty$  will denote the operator norm.

We let  $\mathbb{C}[X_1, \dots, X_n]$  be the free  $*$ -algebra on  $n$ -generators  $X_1, \dots, X_n$ . We will call elements of  $\mathbb{C}[X_1, \dots, X_n]$   $*$ -polynomials in  $n$  indeterminates. For a  $*$ -algebra  $A$ , for elements  $a_1, \dots, a_n \in A$ , and  $P \in \mathbb{C}[X_1, \dots, X_n]$  we use  $P(a_1, \dots, a_n)$  for the image of  $P$  under the unique  $*$ -homomorphism  $\mathbb{C}[X_1, \dots, X_n] \rightarrow A$  sending  $X_j$  to  $a_j$ .

**Definition 2.3.** Let  $(A, \tau)$  be a tracial  $*$ -algebra. An *embedding sequence* is a sequence  $\Sigma = (\sigma_i: A \rightarrow M_{d_i}(\mathbb{C}))$  such that

$$\sup_i \|\sigma_i(a)\|_\infty < \infty, \text{ for all } a \in A$$

$$\|P(\sigma_i(a_1), \dots, \sigma_i(a_n)) - \sigma_i(P(a_1, \dots, a_n))\|_2 \rightarrow 0, \text{ for all } a_1, \dots, a_n \in A \text{ and all } P \in \mathbb{C}[X_1, \dots, X_n]$$

$$\text{tr}(\sigma_i(a)) \rightarrow \tau(a) \text{ for all } a \in A.$$

We will frequently use the following fact: if  $x_1, \dots, x_n \in A$  and  $P \in \mathbb{C}[X_1, \dots, X_n]$  then

$$(3) \quad \|P(\sigma_i(x_1), \dots, \sigma_i(x_n))\|_2 \rightarrow \|P(x_1, \dots, x_n)\|_2.$$

To prove this, first note that as

$$\|P(\sigma_i(x_1), \dots, \sigma_i(x_n)) - \sigma_i(P(x_1, \dots, x_n))\|_2 \rightarrow 0,$$

it suffices to handle the case  $n = 1$  and  $P(X) = X$ . In this case,

$$\|\sigma_i(x)\|_2^2 = \text{tr}(\sigma_i(x)^* \sigma_i(x))$$

and since  $\|\sigma_i(x)^* \sigma_i(x) - \sigma_i(x^*x)\|_2 \rightarrow 0$  we have

$$|\text{tr}(\sigma_i(x)^* \sigma_i(x)) - \text{tr}(\sigma_i(x^*x))| \rightarrow 0.$$

Since

$$\text{tr}(\sigma_i(x^*x)) \rightarrow \|x\|_2^2,$$

we have proved (3).

The proof of the next two propositions will be left to the reader.

**Proposition 2.4.** Let  $\Gamma$  be a countable discrete sofic group with sofic approximation  $\Sigma = (\sigma_i: \Gamma \rightarrow S_{d_i})$ . Extend  $\Sigma$  to maps  $\sigma_i: \mathbb{C}(\Gamma) \rightarrow M_{d_i}(\mathbb{C})$  linearly. Then  $\Sigma$  is an embedding sequence of  $(\mathbb{C}(\Gamma), \tau)$ .

**Proposition 2.5.** Let  $(A, \tau)$  be a tracial  $*$ -algebra and  $\Sigma = (\sigma_i: A \rightarrow M_{d_i}(\mathbb{C}))$  be an embedding sequence. If  $\Sigma' = (\sigma'_i: A \rightarrow M_{d_i}(\mathbb{C}))$  is another sequence of functions so that

$$\sup_i \|\sigma'_i(a)\|_\infty < \infty, \text{ for all } a \in A,$$

$$\|\sigma_i(a) - \sigma'_i(a)\|_2 \rightarrow 0, \text{ for all } a \in A,$$

then  $\Sigma'$  is an embedding sequence.



We will in fact need to extend our sofic approximation to the group von Neumann algebra. For this, we use the following.

**Lemma 2.6** (Lemma 5.5 in [20]). *Let  $\Gamma$  be a countable discrete group. Then any embedding sequence for  $\mathbb{C}(\Gamma)$  extends to one for  $L(\Gamma)$ .*

We will use the preceding lemma when  $\Gamma$  is sofic, in combination with Proposition 2.4. We will often need the following volume-packing estimate.

**Lemma 2.7.** *Let  $n \in \mathbb{N}$ , and  $p \in B(\ell^2(n, u_n))$  an orthogonal projection. For any  $M, \varepsilon > 0$  we have*

$$S_\varepsilon(Mp \text{Ball}(\ell^2(n, u_n))) \leq \left( \frac{3M + \varepsilon}{3} \right)^{2 \text{tr}(p)n}.$$

*Proof.* Let  $A \subseteq Mp \text{Ball}(\ell^2(n, u_n))$  be a maximal  $\varepsilon$ -separated subset of  $Mp \text{Ball}(\ell^2(n, u_n))$ . Thus  $A$  is  $\varepsilon$ -dense. Further,

$$(M + \varepsilon/3)p \text{Ball}(\ell^2(n, u_n)) \supseteq \bigcup_{x \in A} x + (\varepsilon/3)p \text{Ball}(\ell^2(n, u_n)),$$

and the right-hand side is a disjoint union. By linear algebra, the real dimension of the image of  $p$  is  $2 \text{tr}(p)n$ . Thus computing volumes:

$$(M + \varepsilon/3)^{2 \text{tr}(p)n} \text{vol}(p \text{Ball}(\ell^2(n, u_n))) \geq |A|(\varepsilon/3)^{2 \text{tr}(p)n} \text{vol}(p \text{Ball}(\ell^2(n, u_n))).$$

Thus

$$|A| \leq \left( \frac{3M + \varepsilon}{\varepsilon} \right)^{2 \text{tr}(p)n}.$$

□

### 3. DEFINITION OF ENTROPY IN THE PRESENCE OF A POLISH MODEL

Our definition will follow the ideas in [23] and will use *dynamically generating* pseudometrics. We need to state the definition so that it works for actions on Polish spaces. We will need to assume that the pseudometrics are bounded as this is no longer automatic in the noncompact case.

**Definition 3.1.** Let  $\Gamma$  be a countable discrete group and  $X$  a Polish space with  $\Gamma \curvearrowright X$  by homeomorphisms. A bounded continuous pseudometric  $\Delta$  on  $X$  is said to be *dynamically generating* if for any  $x \in X$  and any open neighborhood  $U$  of  $x$  in  $X$ , there is a  $\delta > 0$  and a  $F \subseteq \Gamma$  finite so that if  $x \in X$  and  $\max_{g \in F} d(gx, gy) < \delta$ , then  $y \in U$ .

Recall that if  $X$  is compact, and  $\Gamma$  is a countable discrete group acting on  $X$  by homeomorphisms and  $\Delta$  is a continuous pseudometric on  $X$ , then  $\Delta$  is said to be *dynamically generating* if

$$\sup_{g \in \Gamma} \Delta(gx, gy) > 0 \text{ whenever } x \neq y \text{ in } X$$

(see e.g. [29] Section 4). The fact that this is equivalent to our definition is an easy exercise using compactness of  $X$ .

When  $X$  is Polish, as we shall see in the proof of Lemma 3.9 it really is necessary to require the existence of  $U, \delta, F$  as in the preceding definition instead of just

$$\sup_{g \in \Gamma} \Delta(gx, gy) > 0 \text{ whenever } x \neq y \text{ in } X.$$

One way to realize why this is the correct definition is as follows: let  $(X, \Delta, \Gamma)$  be as in the preceding definition and let  $Y$  be  $X$  modded out by the equivalence relation of  $a \sim b$  if  $\Delta(a, b) = 0$ . Give  $Y$  the metric

$$\overline{\Delta}([a], [b]) = \Delta(a, b).$$

Consider the continuous map

$$\Phi: X \rightarrow Y^\Gamma$$

given by

$$\Phi(x)(g) = [gx],$$

where we use  $[a]$  for the equivalence class of  $a \in X$ . The existence of  $U, \delta$  as in the definition is the *precise* requirement one needs to guarantee that  $\Phi$  is a homeomorphism onto its image. This will be explicitly proven in Lemma 3.9.

If  $(X, \Delta)$  is a pseudometric space, we let  $\Delta_2$  be the pseudometric on  $X^n$  defined by

$$\Delta_2(x, y)^2 = \frac{1}{n} \sum_{j=1}^n \Delta(x(j), y(j))^2.$$

**Definition 3.2.** Let  $\Gamma$  be a countable discrete group and  $X$  a Polish space with  $\Gamma \curvearrowright X$  by homeomorphisms. Let  $\Delta$  be a bounded pseudometric on  $X$ . For a function  $\sigma: \Gamma \rightarrow S_d$ , for some  $d \in \mathbb{N}$ , a finite  $F \subseteq \Gamma$ , and a  $\delta > 0$  we let  $\text{Map}(\Delta, F, \delta, \sigma)$  be all functions  $\phi: \{1, \dots, d\} \rightarrow X$  so that

$$\max_{g \in F} \Delta_2(\phi \circ \sigma(g), g\phi) < \delta,$$

We caution the reader that though we shall typically require  $X$  to be Polish, we will not require our pseudometrics to be complete. We will typically only need to care about the topological consequences of being Polish and not any metric properties. Note that  $\text{Map}(\Delta, F, \delta, \sigma)$  does not account for the measure-theoretic structure of  $X$ . Given a Polish space  $X$ , a finite  $L \subseteq C_b(X)$ , a  $\delta > 0$ , and  $\mu \in \text{Prob}(X)$  let

$$U_{L, \delta}(\mu) = \bigcap_{f \in L} \left\{ \nu \in \text{Prob}(X) : \left| \int f d\nu - \int f d\mu \right| < \delta \right\}.$$

Then  $U_{L, \delta}(\mu)$  form a basis of neighborhoods of  $\mu$  for the weak topology. Here  $C_b(X)$  is the space of bounded continuous functions on  $X$ . Recall that  $u_d$  denotes the uniform probability measure on  $\{1, \dots, d\}$ .

**Definition 3.3.** Suppose that  $\mu$  is a  $\Gamma$ -invariant Borel probability measure on  $X$ . For  $F \subseteq \Gamma$  finite,  $\delta > 0$  and  $L \subseteq C_b(X)$  finite, and  $\sigma: \Gamma \rightarrow S_d$  for some  $d \in \mathbb{N}$  we let  $\text{Map}_\mu(\Delta, F, \delta, L, \sigma)$  be the set of all  $\phi \in \text{Map}(\Delta, F, \delta, \sigma)$  so that

$$\phi_*(u_d) \in U_{L, \delta}(\mu).$$

for all  $f \in L$ .

**Definition 3.4.** Let  $\Gamma$  be a countable discrete sofic group with sofic approximation  $\Sigma = (\sigma_i: \Gamma \rightarrow S_{d_i})$ . Let  $X$  be a Polish space with  $\Gamma \curvearrowright X$  by homeomorphisms and  $\mu$  a  $\Gamma$ -invariant, Borel, probability measure on  $X$ . We define the entropy of  $\Gamma \curvearrowright (X, \mu)$  by

$$h_{\Sigma, \mu}(\Delta, \varepsilon, F, \delta, L) = \limsup_{i \rightarrow \infty} \frac{1}{d_i} \log N_\varepsilon(\text{Map}_\mu(\Delta, F, \delta, L, \sigma_i), \Delta_2)$$

$$h_{\Sigma, \mu}(\Delta, \varepsilon) = \inf_{\substack{F \subseteq \Gamma \text{ finite,} \\ \delta > 0, \\ L \subseteq C_b(X) \text{ finite}}} h_{\Sigma, \mu}(\Delta, F, \delta, L, \varepsilon)$$

$$h_{\Sigma, \mu}(\Delta) = \sup_{\varepsilon > 0} h_{\Sigma, \mu}(\Delta, \varepsilon).$$

By (2) we know that  $h_{\Sigma, \mu}(\Delta)$  is unchanged if we replace  $N_\varepsilon$  with  $S_\varepsilon$ . Because we use  $N_\varepsilon$  instead of  $S_\varepsilon$ , we have  $h_{\Sigma, \mu}(\Delta, \varepsilon, F, \delta, L) \leq h_{\Sigma, \mu}(\Delta, \varepsilon, F', \delta', L')$  if  $F \supseteq F', \delta \leq \delta', L \supseteq L'$ .

The reader may be concerned about finiteness of the expression  $N_\varepsilon(\text{Map}_\mu(\Delta, F, \delta, L, \sigma_i), \Delta_2)$ , since  $X$  is not compact. We note that if  $\varepsilon > 0$  is given, there is a finite  $L \subseteq C_b(X)$ , and  $\delta > 0$  so that for any finite  $F \subseteq \Gamma$  we have

$$N_\varepsilon(\text{Map}_\mu(\Delta, F, \delta, L, \sigma_i), \Delta_2) < \infty.$$

To see this, note that by Prokhorov's Theorem we may choose a  $K \subseteq X$  compact so that

$$\mu(K) \geq 1 - \varepsilon.$$

It is not hard to see that if  $L \subseteq C_b(X)$  is sufficiently large and  $\delta > 0$  is sufficiently small, then for any  $\phi \in \text{Map}_\mu(\Delta, F, \delta, L, \sigma_i)$ ,

$$\frac{1}{d} |\{j : \phi(j) \notin K\}| = \phi_*(u_d)(K^c) \leq 2\varepsilon.$$

Suppose that  $S$  is a finite  $\varepsilon$ -dense subset of  $K$ . Let  $M$  be the diameter of  $(X, \Delta)$  and fix  $x_0 \in X$ . If we set  $A = \phi^{-1}(K)$ , then we find that  $u_{d_i}(A) \geq 1 - \varepsilon$ . We can find a  $\psi \in S^A$  so that if  $\tilde{\psi}: \{1, \dots, d_i\} \rightarrow X$  is defined by

$$\tilde{\psi}(j) = \begin{cases} \psi(j), & \text{if } j \in A, \\ x_0, & \text{if } j \in \{1, \dots, d_i\} \setminus A. \end{cases}$$

Then

$$\Delta_2(\phi, \tilde{\psi})^2 \leq 2M^2\varepsilon + \varepsilon^2.$$

Thus

$$N_{2(2M^2\varepsilon + \varepsilon^2)^{1/2}}(\text{Map}_\mu(\Delta, F, \delta, L, \sigma_i), \Delta_2) \leq S_{(2M^2\varepsilon + \varepsilon^2)^{1/2}}(\text{Map}_\mu(\Delta, F, \delta, L, \sigma_i), \Delta_2) \leq S_\varepsilon(K) < \infty.$$

The main goal of this section is to show that  $h_{\Sigma, \mu}(\Delta)$  is the same as the measure entropy of  $\Gamma \curvearrowright (X, \mu)$  as defined by Bowen and extended by Kerr-Li. Throughout we use  $h_{\Sigma, \mu}(X, \Gamma)$  for sofic measure entropy as defined by [5], [25]. We will use the formulation of sofic entropy in terms of partitions due to Kerr in [22]. However, we will use the terminology of observables as Bowen did in [6].

**Definition 3.5.** Let  $(X, \mathcal{M}, \mu)$  be a standard probability space. Let  $\mathcal{S}$  be a subalgebra of  $\mathcal{M}$  (here  $\mathcal{S}$  is not necessarily a  $\sigma$ -algebra). A *finite  $\mathcal{S}$ -measurable observable* is a measurable map  $\alpha: X \rightarrow A$  where  $A$  is a finite set and  $\alpha^{-1}(\{a\}) \in \mathcal{S}$  for all  $a \in A$ . If  $\mathcal{S} = \mathcal{M}$  we simply call  $\alpha$  a *finite observable*. Another finite  $\mathcal{S}$ -measurable observable  $\beta: X \rightarrow B$  is said to *refine*  $\alpha$ , written  $\alpha \leq \beta$ , if there is a  $\pi: B \rightarrow A$  so that  $\pi(\beta(x)) = \alpha(x)$  for almost every  $x \in X$ . If  $\Gamma$  is a countable discrete group and  $\Gamma \curvearrowright (X, \mathcal{M}, \mu)$  by measure-preserving transformations we say that  $\mathcal{S}$  is generating if  $\mathcal{M}$  is the  $\sigma$ -algebra generated by  $\{gA : A \in \mathcal{S}\}$  (up to sets of measure zero).

For the next definition we need to set up some notation. Given a standard probability space  $(X, \mathcal{M}, \mu)$ , a countable discrete group with  $\Gamma \curvearrowright (X, \mathcal{M}, \mu)$  by measure-preserving transformations, a finite observable  $\alpha: X \rightarrow A$  and  $F \subseteq \Gamma$  finite, we let

$$\tilde{\alpha}^F: X \rightarrow A^F$$

be defined by

$$\tilde{\alpha}^F(x)(g) = \alpha(g^{-1}x).$$

**Definition 3.6.** Let  $\Gamma$  be a countable discrete group and  $\sigma \in S_d^\Gamma$  for some  $d \in \mathbb{N}$ . Let  $(X, \mathcal{M}, \mu)$  be a standard probability space and let  $\mathcal{S} \subseteq \mathcal{M}$  be a subalgebra. Let  $\alpha: X \rightarrow A$  be a finite  $\mathcal{S}$ -measurable-observable. Given  $F \subseteq \Gamma$  finite, and  $\delta > 0$ , we let  $\text{AP}(\alpha, F, \delta, \sigma)$  be the set of all  $\phi: \{1, \dots, d\} \rightarrow A^F$  so that

$$\sum_{a \in A^F} |u_d((\phi^{-1}(\{a\})) - \mu((\tilde{\alpha}^F)^{-1}(\{a\})))| < \delta.$$

$$u_{d_i}(\{j : \phi(j)(g) = \phi(\sigma_i(g)^{-1}(j))(e)\}) < \delta, \text{ for all } g \in F.$$

We now give Kerr's definition of sofic measure entropy in [22].

**Definition 3.7.** Let  $\Gamma$  be a countable discrete sofic group with sofic approximation  $\Sigma = (\sigma_i: \Gamma \rightarrow S_{d_i})$ . Let  $(X, \mathcal{M}, \mu)$  be a standard probability space and  $\Gamma \curvearrowright (X, \mathcal{M}, \mu)$  by measure-preserving transformations. Let  $\mathcal{S}$  be a subalgebra of  $\mathcal{M}$ . Let  $\alpha: X \rightarrow A$  be a finite  $\mathcal{S}$ -measurable observable, and let  $\beta: X \rightarrow B$  refine  $\alpha$ , and  $\pi: B \rightarrow A$  as in the definition of  $\alpha \leq \beta$ . For any  $F \subseteq \Gamma$  finite, we use

$$\pi^*: B^F \rightarrow A$$

for

$$\pi^*(b) = \pi(b(e)).$$

We set

$$h_{\Sigma, \mu}(\alpha; \beta, F, \delta) = \limsup_{i \rightarrow \infty} \frac{1}{d_i} \log |(\pi^*)^{d_i}(\text{AP}(\beta, F, \delta, \sigma_i))|$$

$$h_{\Sigma, \mu}(\alpha; \beta) = \inf_{\substack{F \subseteq \Gamma \text{ finite,} \\ \delta > 0}} h_{\Sigma, \mu}(\alpha; \beta, F, \delta).$$

We then set

$$h_{\Sigma, \mu}(\alpha; \mathcal{S}) = \inf_{\alpha \leq \beta} h_{\Sigma, \mu}(\alpha; \beta)$$

$$h_{\Sigma,\mu}(\mathcal{S}) = \sup_{\alpha} h_{\Sigma,\mu}(\alpha; \beta)$$

where the last infimum and supremum are over all  $\mathcal{S}$ -measurable observables.

We need the following result of Kerr.

**Theorem 3.8.** *Let  $\Gamma$  be a countable discrete sofic group with sofic approximation  $\Sigma$ . Let  $(X, \mathcal{M}, \mu)$  be a standard probability space with  $\Gamma \curvearrowright (X, \mathcal{M}, \mu)$  by measure-preserving transformations. Let  $\mathcal{S} \subseteq \mathcal{M}$  be a generating subalgebra. Then*

$$h_{\Sigma,\mu}(\mathcal{S}) = h_{\Sigma,\mu}(X, \Gamma).$$

Additionally, one can show that

$$h_{\Sigma,\mu}(\alpha; \mathcal{S})$$

is independent of  $\mathcal{S}$  if  $\mathcal{S}$  generates  $\mathcal{M}$  (up to sets of measure zero). In this case, we set

$$h_{\Sigma,\mu}(\alpha; \mathcal{S}) = h_{\Sigma,\mu}(\alpha; X, \Gamma).$$

We now proceed to prove that our definition of sofic entropy with respect to a Polish model recovers measure-theoretic entropy with respect to a sofic approximation. Let us briefly outline the proof. First we show that for any dynamically generating pseudometric  $\Delta$  on  $X$ , there is a compatible metric  $\Delta'$  so that

$$h_{\Sigma,\mu}(\Delta) = h_{\Sigma,\mu}(\Delta')$$

(see Lemma 3.9). Thus we may assume that our metrics are compatible. We then use Kerr's version of measure entropy, using the subalgebra of sets which from the point of view of the measure "appear" to be open and closed (in a sense to be made precise later). We then show that both the topological version and the observable version of microstates produce roughly the same space (see Lemma 3.11). The essential fact for proving the last step will be tightness of a single probability measure on a Polish space. The theorem will follow without too much difficulty from these preliminary lemmas. The following proof is a minor modification of the argument in Lemma 4.4 of [30], as well as Lemma 6.12 of [7]. We have decided to include the proof to alleviate any concerns that may arise from working in the noncompact case, as well as address the necessary modifications that occur in the definition of a dynamically generating pseudometric in the Polish case.

**Lemma 3.9.** *Let  $\Gamma$  be a countable discrete sofic group with sofic approximation  $\Sigma = (\sigma_i: \Gamma \rightarrow S_{d_i})$ . Let  $X$  be a Polish space with  $\Gamma \curvearrowright X$  by homeomorphisms and  $\mu$  a  $\Gamma$ -invariant probability measure. Given a dynamically generating pseudometric  $\Delta$  on  $X$ , there is a bounded compatible metric  $\Delta'$  on  $X$  so that*

$$h_{\Sigma,\mu}(\Delta) = h_{\Sigma,\mu}(\Delta').$$

*Proof.* Let  $M$  be the diameter of  $(X, \Delta)$ . Since  $\Gamma$  is countable, we may find positive real numbers  $\{\alpha_g : g \in \Gamma\}$  with  $\alpha_e \geq 1/2$  and

$$\sum_g \alpha_g = 1.$$

Set

$$\Delta'(x, y) = \sum_{g \in \Gamma} \alpha_g \Delta(gx, gy).$$

We prove the lemma with this  $\Delta'$ . The lemma is proved in several steps.

*Step 1: We show that  $\Delta'$  is a compatible metric.* For this let  $Y$  be  $X$  modded out by the equivalence relation  $a \sim b$  if  $\Delta(a, b) = 0$ . For  $a \in X$ , let  $[a]$  be the equivalence class of  $a$ . Make  $Y$  a metric space with metric  $\overline{\Delta}$  given by

$$\overline{\Delta}([a], [b]) = \Delta(a, b).$$

Observe that  $\overline{\Delta}$  is well-defined because  $\Delta$  satisfies the triangle inequality. Then  $Y^\Gamma$  has a compatible metric given by

$$\Delta_\Gamma(a, b) = \sum_{g \in \Gamma} \alpha_g \overline{\Delta}(a(g), b(g)).$$

Moreover we have an injective map  $\Phi: X \rightarrow Y^\Gamma$  by

$$\Phi(x)(g) = [gx],$$

as

$$\Delta_\Gamma(\Phi(x), \Phi(y)) = \Delta'(x, y),$$

it is enough to show that  $\Phi$  is a homeomorphism onto its image. It is clear that  $\Phi$  is continuous. Suppose that  $x \in X$  and  $U$  is a neighborhood of  $x \in X$ . By the definition of dynamically generating, we may choose a finite  $F \subseteq \Gamma$  and a  $\delta > 0$  so that if  $y \in X$  and

$$\max_{g \in F} \Delta(gx, gy) < \delta,$$

then  $y \in U$ . If we let

$$V = \{y \in Y^\Gamma : \overline{\Delta}([gx], y(g)) < \delta \text{ for all } g \in F\},$$

then  $V$  is a neighborhood of  $\Phi(x)$  in  $Y^\Gamma$  and if  $y \in X$  has  $\Phi(y) \in V$ , then  $y \in U$ . Thus  $\Phi$  is a homeomorphism onto its image.

*Step 2: We show that  $h_{\Sigma, \mu}(\Delta') \leq h_{\Sigma, \mu}(\Delta)$ .* Let  $\varepsilon > 0$  and choose a finite  $E \subseteq \Gamma$  sufficiently large so that

$$\sum_{g \in \Gamma \setminus E} \alpha_g < \varepsilon.$$

Let  $F \subseteq \Gamma$  finite,  $\delta > 0$  and  $L \subseteq C_b(X)$  finite be given. We will assume that  $F \supseteq E$ . We will choose  $\delta$  sufficiently small in a manner depending upon  $\varepsilon$  to be determined later. Since  $\alpha_e \geq 1/2$  we have

$$\Delta(x, y) \leq 2\Delta'(x, y).$$

So by Minkowski's inequality

$$\text{Map}_\mu(\Delta', F, \delta, L, \sigma_i) \subseteq \text{Map}_\mu(\Delta, F, 4\delta, L, \sigma_i).$$

For  $\phi, \psi \in \text{Map}_\mu(\Delta, F, 2\delta, L, \sigma_i)$  we have by Minkowski's inequality,

$$\begin{aligned} \Delta'_2(\phi, \psi) &\leq \sum_{g \in \Gamma} \alpha_g \Delta_2(g\phi, g\psi) \\ &\leq \varepsilon M + \sum_{g \in E} \alpha_g \Delta_2(g\phi, g\psi) \\ &\leq \varepsilon M + 2\delta + \sum_{g \in E} \alpha_g \Delta_2(\phi \circ \sigma_i(g), \Delta \circ \sigma_i(g)) \\ &\leq \varepsilon M + 2\delta + \Delta_2(\phi, \psi) \end{aligned}$$

where in the last line we use that  $\Delta_2(\phi \circ \sigma_i(g), \psi \circ \sigma_i(g)) = \Delta_2(\phi, \psi)$ . Thus for all sufficiently small  $\delta$

$$S_{3\varepsilon(M+2)}(\text{Map}_\mu(\Delta', F, \delta, L, \sigma_i), \Delta'_2) \leq S_\varepsilon(\text{Map}_\mu(\Delta, F, 4\delta, L, \sigma_i)).$$

Thus by (2) we have

$$h_{\Sigma, \mu}(\Delta', 6\varepsilon(M+2)) \leq h_{\Sigma, \mu}(\Delta, \varepsilon, F, 4\delta, L).$$

As  $h_{\Sigma, \mu}(\Delta, \varepsilon, F, 4\delta, L)$  is monotone in  $(F, 4\delta, L)$  we let  $F, L$  increase to  $\Gamma, C_b(X)$  and take  $\delta \rightarrow 0$  to find that

$$h_{\Sigma, \mu}(\Delta', 6\varepsilon(M+2)) \leq h_{\Sigma, \mu}(\Delta', \varepsilon) \leq h_{\Sigma, \mu}(\Delta).$$

Now letting  $\varepsilon \rightarrow 0$  completes the proof of Step 2.

*Step 3: We show that  $h_{\Sigma, \mu}(\Delta) \leq h_{\Sigma, \mu}(\Delta')$ .* Let  $\varepsilon > 0$ . Suppose we are given finite  $F' \subseteq \Gamma, L' \subseteq C_b(X)$  and  $\delta' > 0$ . Set  $L = L'$  and let  $F \subseteq \Gamma$  be a sufficiently large finite set depending upon  $F', \delta'$  in a manner to be determined later and set  $\delta = \delta'$ . Choose a finite  $E \subseteq \Gamma$  so that

$$\sum_{g \in \Gamma \setminus E} \alpha_g < \delta'.$$

If  $\phi \in \text{Map}_\mu(\Gamma, F, \delta, L, \sigma_i)$  and  $h \in F'$ , we have by Minkowski's inequality

$$\begin{aligned}
\Delta'_2(h\phi, \phi \circ \sigma_i(h)) &\leq \sum_{g \in \Gamma} \alpha_g \Delta_2(gh\phi, g\phi \circ \sigma_i(h)) \\
&\leq \delta' + \sum_{g \in F} \alpha_g \Delta_2(gh\phi, \phi \circ \sigma_i(gh)) + \sum_{g \in F} \alpha_g \Delta_2(\phi \circ \sigma_i(gh), \phi \circ \sigma_i(g)\sigma_i(h)) \\
&\quad + \sum_{g \in F} \alpha_g \Delta_2(\phi \circ \sigma_i(g)\sigma_i(h), g\phi \circ \sigma_i(h)) \\
&= \delta' + \sum_{g \in F} \alpha_g \Delta_2(gh\phi, \phi \circ \sigma_i(gh)) + \sum_{g \in F} \alpha_g \Delta_2(\phi \circ \sigma_i(gh), \phi \circ \sigma_i(g)\sigma_i(h)) \\
&\quad + \sum_{g \in F} \alpha_g \Delta_2(\phi \circ \sigma_i(g), g\phi) \\
&< \delta' + 2\delta + \sum_{g \in E} \alpha_g \Delta_2(\phi \circ \sigma_i(g)\sigma_i(h), \phi \circ \sigma_i(gh)),
\end{aligned}$$

if we force  $F \supseteq EF' \cup E$ . Set  $\delta = \delta'$ , as

$$\Delta_2(\phi \circ \sigma_i(g)\sigma_i(h), \phi \circ \sigma_i(gh))^2 \leq Mu_{d_i}(\{1 \leq j \leq d_i : \sigma_i(g)\sigma_i(h)(j) \neq \sigma_i(gh)(j)\}) \rightarrow 0$$

for  $g, h \in \Gamma$  we find that

$$\text{Map}_\mu(\Delta, F, \delta, L, \sigma_i) \subseteq \text{Map}_\mu(\Delta', F', 4\delta', \sigma_i),$$

for all large  $i$ . As

$$\Delta_2(\phi, \psi) \leq 2\Delta'_2(\phi, \psi),$$

we find that

$$h_{\Sigma, \mu}(\Delta, 6\varepsilon) \leq h_{\Sigma, \mu}(\Delta, 6\varepsilon, F, \delta, L) \leq h_{\Sigma, \mu}(\Delta', \varepsilon, F', \delta', L').$$

Taking the infimum over  $F', \delta', L'$  we find that

$$h_{\Sigma, \mu}(\Delta, 6\varepsilon) \leq h_{\Sigma, \mu}(\Delta', \varepsilon) \leq h_{\Sigma, \mu}(\Delta').$$

Letting  $\varepsilon \rightarrow 0$  proves Step 3. □

To prove the theorem, we need to single out a nice subalgebra of measurable sets. Let  $X$  be a Polish space and  $\mu$  a Borel probability measure on  $X$ . We let  $\text{CO}_\mu$  be the set of Borel sets  $E$  so that

$$\mu(\text{int } E) = \mu(\overline{E}).$$

Note that  $\text{CO}_\mu$  is an algebra of sets. These sets are often called continuity sets in the literature.

For the next lemma we need some notation. Given a metric space  $(X, \Delta)$ ,  $E \subseteq X$  and  $\varepsilon > 0$  we let

$$E_\varepsilon = \{x \in E : \Delta(x, E^c) \geq \varepsilon\}$$

$$\mathcal{O}_\varepsilon(E) = \{x \in X : \Delta(x, E) < \varepsilon\}.$$

**Lemma 3.10.** *Let  $X$  be a Polish space and  $\mu$  a Borel probability measure on  $X$ . Let  $\Delta$  be a compatible metric on  $X$ . Given  $E \in \text{CO}_\mu$ , and  $\eta > 0$  there is a neighborhood  $U$  of  $\mu$  in the weak topology and a  $\kappa > 0$  so that if  $\nu \in U \cap \text{Prob}(X)$  then*

$$|\nu(E) - \mu(E)| < \eta$$

$$\nu(\mathcal{O}_\kappa(E) \setminus E_\kappa) < \eta.$$

*Proof.* It is a consequence of the Portmanteau Theorem that a sequence of probability measure  $\mu_n$  on  $X$  converges weakly to  $\mu$  if and only if  $\mu_n(A) \rightarrow \mu(A)$  for every continuity set  $A$  of  $\mu$ . Thus we can choose a neighborhood  $U_1$  of  $\mu$  so that

$$|\nu(E) - \mu(E)| < \eta$$

for all  $\nu \in U_1$ . To obtain the second estimate, again by the Portmanteau Theorem we can choose a neighborhood  $U_2$  of  $\mu$  so that

$$\nu(\overline{\mathcal{O}_{2\kappa}(E) \setminus E_{2\kappa}}) \leq \mu(\overline{\mathcal{O}_{2\kappa}(E) \setminus E_{2\kappa}}) + \frac{\eta}{2}$$

for all  $\nu \in U_2$ . We have  $\mathcal{O}_\kappa(E) \setminus E_\kappa \subseteq \overline{\mathcal{O}_{2\kappa}(E) \setminus E_{2\kappa}}$ , so

$$\nu(\mathcal{O}_\kappa(E) \setminus E_\kappa) \leq \mu(\overline{\mathcal{O}_{2\kappa}(E) \setminus E_{2\kappa}}) + \frac{\eta}{2}.$$

Since  $E$  is a continuity set of  $\mu$  we can choose  $\kappa$  small enough so that

$$\mu(\overline{\mathcal{O}_{2\kappa}(E) \setminus E_{2\kappa}}) < \frac{\eta}{2}.$$

The lemma is now completed by setting  $U = U_1 \cap U_2$ . □

Given  $\sigma: \Gamma \rightarrow S_d$  for some  $d \in \mathbb{N}$ , and  $\phi \in A^d$  we shall define  $\phi^F: \{1, \dots, d\} \rightarrow A^F$  by

$$\phi_\sigma^F(j)(g) = \phi(\sigma(g)^{-1}(j)), \text{ for } g \in F$$

**Lemma 3.11.** *Let  $\Gamma$  be a countable discrete sofic group with sofic approximation  $\Sigma = (\sigma_i: \Gamma \rightarrow S_{d_i})$ . Let  $X$  be a Polish space with  $\Gamma \curvearrowright X$  by homeomorphisms and  $\mu$  a  $\Gamma$ -invariant Borel probability measure on  $X$ . Let  $\Delta$  be a bounded compatible metric on  $X$ .*

(i): *Let  $\beta: X \rightarrow B$  a finite  $\text{CO}_\mu$ -measurable observable. Given  $F \subseteq \Gamma$  finite,  $\delta > 0$  there are finite  $F' \subseteq \Gamma, L' \subseteq C_b(X)$ , and  $\delta' > 0$  so that if  $\phi \in \text{Map}_\mu(\Delta, F', \delta', L', \sigma_i)$  then  $(\beta \circ \phi)_{\sigma_i}^F \in \text{AP}(\beta, F, \delta, \sigma_i)$ .*

(ii): *Given finite  $F' \subseteq \Gamma, L' \subseteq C_b(X)$  and  $\delta' > 0$ , there are finite  $F \subseteq \Gamma$  and a  $\delta > 0$  and a finite  $\text{CO}_\mu$ -measurable observable  $\beta: X \rightarrow B$  so that if  $\tilde{\beta}^F \circ \phi \in \text{AP}(\beta, F, \delta, \sigma_i)$  then  $\phi \in \text{Map}_\mu(\Delta, F', \delta', L', \sigma_i)$ .*

*Proof.* (i): Let  $\eta > 0$  be sufficiently small depending upon  $\delta$  in a manner to be determined later. By the preceding lemma, we may find a  $L' \subseteq C_b(X)$  finite, and a  $\delta' > 0$  so that if  $\nu \in \text{Prob}(X)$  and

$$\left| \int_X f d\mu - \int_X f d\nu \right| < \delta' \text{ for all } f \in L'$$

then for all  $E \subseteq F$  and for all  $b \in B^E$ ,

$$|\mu((\tilde{\beta}^E)^{-1}(\{b\})) - \nu((\tilde{\beta}^E)^{-1}(\{b\}))| < \eta,$$

$$\nu(\mathcal{O}_{\sqrt{\delta'}}(\tilde{\beta}^E)^{-1}(\{b\}) \setminus (\tilde{\beta}^E)^{-1}(\{b\})_{\sqrt{\delta'}}) \leq \eta.$$

Set  $F' = F$ . Suppose that  $\phi \in \text{Map}_\mu(\Delta, F', \delta', L', \sigma_i)$ . Note that if we choose  $\eta$  sufficiently small, then we have forced

$$\sum_{b \in B^F} |\mu((\tilde{\beta}^F)^{-1}(\{b\})) - u_{d_i}(((\beta \circ \phi)_{\sigma_i}^F)^{-1}(\{b\}))| < \delta$$

for all  $\phi \in \text{Map}(\Delta, F', \delta', L', \sigma_i)$ . Since  $(\beta \circ \phi)_{\sigma_i}^F(j)(g) = (\beta \circ \phi)^F(\sigma_i(g)^{-1}(j))(e)$ , this means  $(\beta \circ \phi)_{\sigma_i}^F \in \text{AP}(\beta, F, \delta, \sigma_i)$ .

(ii): Let  $M > 0$  be the diameter of  $(X, \Delta)$ . Let  $\kappa > 0$  be sufficiently small depending upon  $F', \delta'$  in a manner to be determined later. Since  $X$  is Polish, Prokhorov's Theorem applied to  $\{\mu\}$  implies that we can find a compact set  $K \subseteq X$  so that

$$\mu(X \setminus K) \leq \kappa.$$

Since  $K$  is compact, we can find points  $x_1, \dots, x_n \in K$ , and numbers  $\delta_j \in (0, \kappa), j = 1, \dots, n$  so that if  $B(x, \alpha)$  is the ball in  $X$  of radius  $\alpha$  with respect to  $\Delta$  then

$$\mu(\overline{B(x_j, \delta_j)} \setminus B(x_j, \delta_j)) = 0,$$

$$\sup_{x, y \in B(x_j, \delta_j)} |f(x) - f(y)| < \delta' \text{ for all } f \in L',$$

$$K \subseteq \bigcup_{j=1}^n B(x_j, \delta_j).$$

Let

$$E = X \setminus \bigcup_{j=1}^n B(x_j, \delta_j)$$

and define  $\beta: X \rightarrow \{0, 1\}^{n+1}$  by

$$\beta(x)(k) = \begin{cases} \chi_{B(x_k, \delta_k)}(x), & \text{if } 1 \leq k \leq n \\ \chi_E(x), & \text{if } k = n+1. \end{cases}$$

Note that  $\beta$  is  $\text{CO}_\mu$ -measurable. Set  $F = F' \cup \{e\} \cup (F')^{-1}$ , and let  $\delta > 0$  be sufficiently small in a manner to depend upon  $L', \delta', F'$  to be determined later. We will assume that

$$\delta < \min\{\delta_j : j = 1, \dots, n\}.$$

Suppose that  $\tilde{\beta}^F \circ \phi \in \text{AP}(\beta, F, \delta, \sigma_i)$ . Given  $g \in F$ , let

$$C_g = \{j : \beta(g\phi(j)) = \beta(\phi(\sigma_i(g)(j)))\}.$$

Since

$$\beta(g\phi(j)) = (\tilde{\beta}^F \circ \phi)(j)(g^{-1}),$$

$$\beta(\phi(\sigma_i(g)(j))) = (\tilde{\beta}^F \circ \phi)(\sigma_i(g)(j))(e),$$

and

$$u_{d_i}(\{j : \sigma_i(g^{-1})^{-1}(j) = \sigma_i(g)(j)\}) \rightarrow_{i \rightarrow \infty} 1$$

we see that for all large  $i$ ,

$$u_{d_i}(C_g) \geq 1 - 2\delta.$$

For  $j \in C_g$ , we necessarily have

$$\Delta(\phi(\sigma_i(g)(j)), g\phi(j)) < \kappa.$$

Thus

$$\Delta_2(\phi \circ \sigma_i(g), g\phi)^2 \leq \kappa^2 + M^2\delta.$$

So if we choose  $\kappa < \delta$ , and then  $\delta$  sufficiently small, we have forced

$$\Delta_2(\phi \circ \sigma_i(g), g\phi) < 2\delta'.$$

We will want to force  $\kappa, \delta'$  to be even smaller later. Using  $\|f\|$  for the uniform norm of  $f \in C_b(X)$ , we have for all  $f \in L'$ ,

$$\begin{aligned} \left| \int_X f d\mu - \int_X f d\phi_*(u_{d_i}) \right| &\leq \kappa\|f\| + u_{d_i}(\phi^{-1}(E))\|f\| + \sum_{j=1}^n \left| \int_{B(x_j, \delta_j)} f d\mu - \int_{B(x_j, \delta_j)} f d\phi_*(u_{d_i}) \right| \\ &\leq \kappa\|f\| + 2\kappa\|f\| + \delta' + \sum_{j=1}^n \left| f(x_j)\mu(B(x_j, \delta_j)) - \frac{1}{d_i} \sum_{k: \phi(k) \in B(x_j, \delta_j)} f(\phi(k)) \right| \\ &\leq \kappa\|f\| + 2\kappa\|f\| + \delta' + \delta' \sum_{j=1}^n \phi_*(u_{d_i})(B(x_j, \delta_j)) \\ &\quad + \sum_{j=1}^n \left| f(x_j)\mu(B(x_j, \delta_j)) - \frac{1}{d_i} \sum_{k: \phi(k) \in B(x_j, \delta_j)} f(x_j) \right| \\ &\leq \kappa\|f\| + 2\kappa\|f\| + 2\delta' + \sum_{j=1}^n |f(x_j)| |\mu(B(x_j, \delta_j)) - \phi_*(u_{d_i})(B(x_j, \delta_j))| \\ &\leq \kappa\|f\| + 2\kappa\|f\| + 2\delta' + n\|f\|\delta. \end{aligned}$$

We may choose  $\kappa < \frac{\delta'}{\|f\|}$ . This forces  $n$  on us, but we may then choose  $\delta$  sufficiently small so that  $\phi \in \text{Map}_\mu(\Delta, F', 6\delta', L', \sigma_i)$ . As  $\delta'$  is arbitrary this completes the proof.  $\square$

We are now ready to show that our definition of entropy in the case of a Polish model agrees with the usual measure entropy.

**Theorem 3.12.** *Let  $\Gamma$  be a countable discrete sofic group with sofic approximation  $\Sigma$ . Let  $X$  be a Polish space with  $\Gamma \curvearrowright X$  by homeomorphisms and  $\mu$  a  $\Gamma$ -invariant, Borel, probability measure on  $X$ . For any dynamically generating pseudometric  $\Delta$  on  $X$  we have*

$$h_{\Sigma, \mu}(\Delta) = h_{\Sigma, \mu}(X, \Gamma).$$



*Proof.* Let  $\Sigma = (\sigma_i : \Gamma \rightarrow S_{d_i})$ . By Lemma 3.9 we may assume that  $\Delta$  is a bounded compatible metric on  $X$ . Let  $M$  be the diameter of  $(X, \rho)$ . We will apply Theorem 3.8 with  $\mathcal{S} = \text{CO}_\mu$ . We leave it as an exercise to show that for all  $x \in X$  we have

$$\mu(B(x, r)) = \mu(\overline{B(x, r)})$$

for all but countably many  $r \in (0, \infty)$ . Thus  $\text{CO}_\mu$  generates the  $\sigma$ -algebra of Borel subsets of  $X$ . We first show that

$$h_{\Sigma, \mu}(\Delta) \leq h_{\Sigma, \mu}(\text{CO}_\mu).$$

Let  $\varepsilon > 0$ . Since  $X$  is Polish, we may apply Prokhorov's Theorem to find a compact  $K \subseteq X$  so that

$$\mu(X \setminus K) < \varepsilon.$$

By compactness of  $K$ , we find  $x_1, \dots, x_n \in K$ , and  $\varepsilon > \delta_1, \dots, \delta_n > 0$  so that

$$K \subseteq \bigcup_{j=1}^n B(x_j, \delta_j),$$

$$B(x_j, \delta_j) \in \text{CO}_\mu.$$

Set

$$E = X \setminus \bigcup_{j=1}^n B(x_j, \delta_j).$$

Define

$$\alpha : X \rightarrow \{0, 1\}^{n+1}$$

by

$$\alpha(x)(k) = \begin{cases} \chi_{B(x_k, \delta_k)}(x), & \text{if } 1 \leq k \leq n \\ \chi_E(x), & \text{if } k = n+1. \end{cases}$$

Let  $\beta : X \rightarrow B$  be any finite  $\text{CO}_\mu$ -measurable observable refining  $\alpha$  and let  $\pi : B \rightarrow A$  be as in the definition of  $\beta \geq \alpha$ . Suppose we are given a finite  $F \subseteq \Gamma$  and a  $\delta > 0$ . By the preceding Lemma, we may find a finite  $F' \subseteq \Gamma, L' \subseteq C_b(X)$ , and a  $\delta' > 0$  so that

$$\beta_{\sigma_i}^F(\text{Map}_\mu(\Delta, F', \delta', L', \sigma_i)) \subseteq \text{AP}(\beta, F, \delta, \sigma_i).$$

By Lemma 3.10, we may assume that  $L'$  is sufficiently large so that

$$\phi_*(u_{d_i})(X \setminus E) \leq 2\varepsilon$$

for all  $\phi \in \text{Map}_\mu(\Delta, F', \delta', L', \sigma_i)$ . Choose elements  $\{\phi_s\}_{s \in S}$  with  $\phi_s \in \text{Map}_\mu(\Delta', F', L', \delta', \sigma_i)$  where  $S$  is some index set so that

$$\{\alpha \circ \phi_s : s \in S\} = \alpha^{d_i}(\text{Map}_\mu(\Delta, F', \delta', L', \sigma_i))$$

and so that

$$\alpha \circ \phi_s \neq \alpha \circ \phi_{s'} \text{ for } s \neq s' \text{ in } S.$$

Then

$$|S| \leq |(\pi^*)^{d_i}(\text{AP}(\beta, F, \delta, \sigma_i))|.$$

Let  $\phi \in \text{Map}_\mu(\Delta, F', \delta', L', \sigma_i)$  and let  $s \in S$  be such that  $\alpha \circ \phi = \alpha \circ \phi_s$ . Then

$$\Delta_2(\phi, \phi_s)^2 \leq 4M^2\varepsilon + \frac{1}{d_i} \sum_{j: \phi(j), \phi_s(j) \notin E} \Delta(\phi(j), \phi_s(j))^2.$$

If  $\phi(j)$  and  $\phi_s(j)$  are not in  $E$ , then the fact that  $\alpha(\phi(j)) = \alpha(\phi_s(j))$  implies  $\Delta(\phi(j), \phi_s(j)) < \varepsilon$ , so

$$\Delta_2(\phi, \phi_s)^2 < 4M^2\varepsilon + \varepsilon^2.$$

Thus by (2),

$$h_{\Sigma, \mu}(\Delta, 2(4M\varepsilon + \varepsilon^2)^{1/2}) \leq h_{\Sigma, \mu}(\alpha; \beta, F, \delta).$$

Taking the infimum over all  $\beta, F, \delta$  we find that

$$h_{\Sigma, \mu}(\Delta, 2(4M\varepsilon + \varepsilon^2)^{1/2}) \leq h_{\Sigma, \mu}(\alpha) \leq h_{\Sigma, \mu}(X, \Gamma),$$

and letting  $\varepsilon \rightarrow 0$  implies that

$$h_{\Sigma, \mu}(\Delta) \leq h_{\Sigma, \mu}(X, \Gamma).$$

For the reverse inequality, let  $\alpha: X \rightarrow A$  be a  $\text{CO}_\mu$ -measurable finite observable. Fix  $\kappa > 0$ , and let  $\kappa' > 0$  depend upon  $\kappa$  in a manner to be determined later. By Lemma 3.10 we may choose  $\eta > 0$  and  $L_0 \subseteq C_b(X)$  finite so that if  $\nu \in \text{Prob}(X)$  and

$$\left| \int_X f d\mu - \int_X f d\nu \right| < \eta,$$

for all  $f \in L_0$  then

$$\begin{aligned} |\nu(\alpha^{-1}(\{a\})) - \mu(\alpha^{-1}(\{a\}))| &< \kappa', \\ \nu(\mathcal{O}_\eta(\alpha^{-1}(\{a\}) \setminus \alpha^{-1}(\{a\})_\eta) &< \kappa'. \end{aligned}$$

Let  $F' \subseteq \Gamma, L' \subseteq C_b(X)$  be given finite sets and  $\delta' > 0$  be given. We may assume that  $L' \supseteq L_0$ . By the preceding Lemma, we may choose a refinement  $\beta: X \rightarrow B$  of  $\alpha$ , a finite  $F \subseteq \Gamma$ , and a  $\delta > 0$  so that if  $\phi \in X^{d_i}$  and  $\tilde{\beta}^F \circ \phi \in \text{AP}(\beta, F, \delta, \sigma_i)$ , then  $\phi \in \text{Map}_\mu(\Delta, F', \delta', L', \sigma_i)$ . Choose  $\pi: B \rightarrow A$  so that  $\alpha = \pi \circ \beta$  and choose a map  $s: B^F \rightarrow X$  so that  $\text{Id} = \tilde{\beta}^F \circ s$ . By construction if  $\phi \in \text{AP}(\beta, F', \delta', \sigma_i)$  we have

$$s \circ \phi \in \text{Map}_\mu(\Delta, F', \delta', L', \sigma_i).$$

Let  $\varepsilon > 0$  be sufficiently small depending upon  $\eta$  to be determined later. Let  $\{\phi_t : t \in T\}$  be such that

$$\begin{aligned} \phi_t &\in \text{AP}(\beta, F, \delta, \sigma_i) \text{ for all } t \in T, \\ \{s \circ \phi_t : t \in T\} &\text{ is } \varepsilon\text{-dense in } \{s \circ \phi : \phi \in \text{AP}(\beta, F, \delta, \sigma_i)\}, \\ \phi_t &\neq \phi_{t'} \text{ for } t \neq t'. \end{aligned}$$

We may (and will) choose  $T$  so that

$$|T| \leq S_{\varepsilon/2}(\text{Map}_\mu(\Delta, F', \delta', L', \sigma_i)) \leq N_{\varepsilon/2}(\text{Map}_\mu(\Delta, F', \delta', L', \sigma_i)).$$

Note that

$$(\pi^*)^{d_i}(\text{AP}(\beta, F, \delta, \sigma_i)) \subseteq \bigcup_{t \in T} \alpha^{d_i}(B_{\Delta_2}(s \circ \phi_t, \varepsilon) \cap \text{Map}_\mu(\Delta, F', \delta', L', \sigma_i)).$$

We thus have to bound  $|\alpha^{d_i}(B_{\Delta_2}(\phi_t, \varepsilon) \cap \text{Map}_\mu(\Delta, F', \delta', L', \sigma_i))|$  from above. Fix  $t \in T$ . Suppose that  $\phi \in \text{Map}_\mu(\Delta, F', \delta', L', \sigma_i)$  and that

$$\Delta_2(\phi, s \circ \phi_t) < \varepsilon.$$

Let

$$C = \bigcup_{a \in A} \{1 \leq j \leq d_i : \phi(j) \in \mathcal{O}_\eta(\alpha^{-1}(\{a\})) \setminus \alpha^{-1}(\{a\})_\eta\} \cup \{1 \leq j \leq d_i : s \circ \phi_t(j) \in \mathcal{O}_\eta(\alpha^{-1}(\{a\})) \setminus \alpha^{-1}(\{a\})_\eta\}.$$

By choosing  $\kappa'$  sufficiently small, we may assume that

$$u_{d_i}(C) \leq \kappa.$$

Let

$$D = \{1 \leq j \leq d_i : \Delta(\phi(j), s \circ \phi_t(j)) \geq \sqrt{\varepsilon}\}.$$

Thus

$$u_{d_i}(D) \leq \sqrt{\varepsilon}.$$

For  $j \in \{1, \dots, d_i\} \setminus (C \cup D)$ , and  $a = \alpha(\phi(j))$ , we have that  $s \circ \phi_t(j) \in \mathcal{O}_{\sqrt{\varepsilon}}(\alpha^{-1}(\{a\}))$ . If we choose  $\sqrt{\varepsilon} < \eta$ , then  $\alpha(s \circ \phi_t(j)) = a$  for all  $j \in \{1, \dots, d_i\} \setminus (C \cup D)$ . Thus we can find  $\mathcal{V} \subseteq \{1, \dots, d_i\}$  with  $u_{d_i}(\mathcal{V}) \geq 1 - \sqrt{\kappa} - \sqrt{\varepsilon}$  and  $\alpha(\phi(j)) = \alpha(s \circ \phi_t(j))$  for  $j \in \mathcal{V}$ . Thus

$$\begin{aligned} |\alpha^{d_i}(B_{\Delta_2}(s \circ \phi_t, \varepsilon) \cap \text{Map}_\mu(\Delta, F', \delta', L', \sigma_i))| &\leq \sum_{\substack{\mathcal{V} \subseteq \{1, \dots, d_i\}, \\ |\mathcal{V}| \leq (\kappa + \sqrt{\varepsilon})d_i}} |A|^{|\mathcal{V}|} \\ &\leq \sum_{l=1}^{\lfloor (\kappa + \sqrt{\varepsilon})d_i \rfloor} \binom{d_i}{l} |A|^l. \end{aligned}$$

If  $\kappa + \sqrt{\varepsilon} < 1/2$  then for all large  $i$  we have

$$\binom{d_i}{l} \leq \binom{d_i}{\lfloor \kappa + \sqrt{\varepsilon} \rfloor d_i}.$$

So by Stirling's Formula the above sum is at most

$$R(\kappa + \sqrt{\varepsilon})d_i \exp(d_i H(\kappa + \sqrt{\varepsilon}))|A|^{\kappa d_i}$$

for some constant  $R > 0$ , where

$$H(t) = -t \log t - (1-t) \log(1-t) \text{ for } 0 \leq t \leq 1.$$

Thus

$$h_{\Sigma, \mu}(\alpha) \leq h_{\Sigma, \mu}(\alpha; \beta, F, \delta, \sigma_i) \leq H(\kappa + \sqrt{\varepsilon}) + \kappa \log |A| + h_{\Sigma, \mu}(\Delta, \varepsilon/2, F', \delta', L').$$

Taking the infimum over all  $F', \delta', L'$  and let  $\kappa \rightarrow 0$  we have

$$h_{\Sigma, \mu}(\alpha) \leq h_{\Sigma, \mu}(\Delta, \varepsilon/2) + H(\sqrt{\varepsilon}).$$

Letting  $\varepsilon \rightarrow 0$  and then taking the supremum over all  $\alpha$  we have

$$h_{\Sigma, \mu}(X, \Gamma) \leq h_{\Sigma, \mu}(\Delta).$$

□

#### 4. SPECTRAL CONSEQUENCES OF POSITIVE ENTROPY

Let  $\Gamma \curvearrowright (X, \mu)$  be a probability measure-preserving action of a countable discrete group. Associated to this action we have a natural representation

$$\rho_{\Gamma \curvearrowright (X, \mu)}: \Gamma \rightarrow U(L^2(X, \mu))$$

by

$$(\rho_{\Gamma \curvearrowright (X, \mu)}(g)f)(x) = f(g^{-1}x).$$

The space  $\mathbb{C}1$  inside  $L^2(X, \mu)$  is clearly  $\Gamma$ -invariant, so we can consider the representation  $\rho_{\Gamma \curvearrowright (X, \mu)}^0$  obtained by restricting  $\rho_{\Gamma \curvearrowright (X, \mu)}$  to  $L^2(X, \mu) \ominus \mathbb{C}1$ . The representation  $\rho_{\Gamma \curvearrowright (X, \mu)}^0$  is called the *Koopman* representation. Properties of a probability measure-preserving action are called *spectral* when they only depend upon the Koopman representation. In this section, we deduce spectral properties of an action from assumptions of positive entropy.

**4.1. Representation Theoretic Preliminaries.** We will need to apply the theory of representations of  $*$ -algebras. For this paper, we will only need unitary representations of groups, but later work will need this generality. For notation, if  $A$  is a  $*$ -algebra and  $\rho: A \rightarrow B(\mathcal{H})$  is a  $*$ -representation and  $E$  is a set we use  $\rho^{\oplus E}$  for the  $*$ -representation of  $A$  on  $\ell^2(E, \mathcal{H})$  given by

$$(\rho^{\oplus E}(a)\xi)(x) = \rho(a)\xi(x).$$

Let us mention how the theory for  $*$ -algebras generalizes that of groups. If  $\Gamma$  is a countable discrete group and  $\rho: \Gamma \rightarrow \mathcal{U}(\mathcal{H})$  is a unitary representation, for  $f = \sum_{g \in \Gamma} f_g g \in \mathbb{C}(\Gamma)$  we define

$$\rho(f) = \sum_{g \in \Gamma} f_g \rho(g).$$

We use the conjugate linear, antimultiplicative map  $*$  on  $\mathbb{C}(\Gamma)$  given by

$$f^* = \sum_{g \in \Gamma} \overline{f_{g^{-1}}} g.$$

Under these operations  $\mathbb{C}(\Gamma)$  is  $*$ -algebra and  $\rho(f^*) = \rho(f)^*$  for any unitary representation of  $\Gamma$ .

If  $A$  is a  $*$ -algebra and  $\rho_j: \Gamma \rightarrow \mathcal{U}(\mathcal{H}_j)$  are two  $*$ -representations we write

$$\text{Hom}_A(\rho_1, \rho_2)$$

for the space of bounded linear  $A$ -equivariant maps  $\mathcal{H}_1 \rightarrow \mathcal{H}_2$ .

**Definition 4.1.** Let  $A$  be a  $*$ -algebra and  $\rho_j: A \rightarrow B(\mathcal{H}_j), j = 1, 2$  be  $*$ -representations. We say that  $\rho_1$  and  $\rho_2$  are *mutually singular*, written  $\rho_1 \perp \rho_2$  if for every pair of nonzero subrepresentations  $\rho'_j$  of  $\rho_j$  we have that  $\rho'_1$  is not isomorphic to  $\rho'_2$ . We say that  $\rho_1$  is absolutely continuous with respect to  $\rho_2$ , and write  $\rho_1 \ll \rho_2$  if  $\rho_1$  is embeddable in  $\rho_2^{\oplus E}$  for some set  $E$ .

The terminology is motivated by measure theory. For intuition, suppose that  $A = C(X)$  and that  $\rho_j: A \rightarrow B(\mathcal{H}_j)$ ,  $j = 1, 2$  are  $*$ -homomorphisms. Then we can find spectral measures (in the sense of [13] IX.1)  $E_j$  on  $X$  so that

$$\rho_j(f) = \int_X f(x) dE_j(x).$$

We leave it as an exercise to the reader to check that  $\rho_1 \perp \rho_2$  if and only if  $E_1 \perp E_2$ , and similarly that  $\rho_1 \ll \rho_2$  if and only if  $E_1 \ll E_2$  (the definitions of absolute continuity and singularity of spectral measures is the same as for usual measures).

We will need the following equivalent conditions on singularity of representations. The following must be well known, but we include a proof for completeness. Throughout the proof, we shall use functional calculus. See [13] Chapter VII, IX for background on functional calculus.

**Proposition 4.2.** *Let  $A$  be a unital  $*$ -algebra and  $\rho_j: A \rightarrow B(\mathcal{H}_j)$  be two unitary representations and suppose that  $\mathcal{H}_1, \mathcal{H}_2$  are separable. The following are equivalent:*

- (i):  $\rho_1 \perp \rho_2$
- (ii):  $\text{Hom}_A(\rho_1, \rho_2) = \{0\}$
- (iii):  $\text{Hom}_A(\rho_2, \rho_1) = \{0\}$
- (iv): *There is a sequence  $x_n \in A$  so that  $\max(\|\rho_1(x_n)\|, \|\rho_2(x_n)\|) \leq 1$  and*

$$\rho_1(x_n^* x_n) \rightarrow \text{Id}_{\mathcal{H}_1} \text{ in the strong operator topology}$$

$$\rho_2(x_n^* x_n) \rightarrow 0 \text{ in the strong operator topology.}$$

*Proof.* The equivalence of (ii) and (iii) is proved by taking adjoints. To prove that (ii) implies (i) suppose that  $\mathcal{K}_j$ ,  $j = 1, 2$  are closed,  $A$ -invariant, linear subspaces of  $\mathcal{H}_j$ ,  $j = 1, 2$  and that  $\phi: \mathcal{K}_1 \rightarrow \mathcal{K}_2$  is an isomorphism. Define  $T: \mathcal{H}_1 \rightarrow \mathcal{H}_2$  by

$$T(\xi) = \Phi(\text{proj}_{\mathcal{K}_1}(\xi)).$$

By (ii) we know that  $T = 0$  which implies that  $\mathcal{K}_j$ ,  $j = 1, 2$  are zero.

To see that (i) implies (ii) suppose that  $T \in \text{Hom}_A(\rho_1, \rho_2)$ . Let  $T = U|T|$  be the polar decomposition (see [13] VIII.3.11). The fact that  $T$  is equivariant implies that  $T^*T$  is equivariant, and hence that  $|T| = (T^*T)^{1/2}$  is, by approximating the square root function by polynomials. Since

$$U = \text{SOT} - \lim_{\varepsilon \rightarrow 0} T(|T| + \varepsilon)^{-1}$$

we see that  $U$  is equivariant. Thus  $U$  gives an  $A$ -equivariant isomorphism  $(\ker T)^\perp \rightarrow \overline{\text{im } T}$ . Since  $\rho_1 \perp \rho_2$  we find that  $(\ker T)^\perp = 0$ , and hence that  $T = 0$ .

To prove that (iv) implies (ii) let  $T \in \text{Hom}_A(\rho_1, \rho_2)$ . Let  $x_n$  be as in (iv). Then, for any  $\xi \in \mathcal{H}_1$

$$T(\xi) = \lim_{n \rightarrow \infty} T(\rho_1(x_n^* x_n)\xi) = \lim_{n \rightarrow \infty} \rho_2(x_n^* x_n)T(\xi) = 0,$$

so  $T = 0$ .

Suppose that (iii) and (ii) hold, we wish to prove (iv). Recall that if  $\mathcal{H}$  is a Hilbert space and  $E \subseteq B(\mathcal{H})$  then  $E'$  denotes the commutant of  $E$ , i.e.  $E' = \{T \in B(\mathcal{H}) : TS = ST \text{ for all } S \in E\}$ . Suppose that

$$T \in (\rho_1 \oplus \rho_2)(A)',$$

Then we can regard  $T$  as a matrix

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$$

where  $T_{ij} \in B(\mathcal{H}_i, \mathcal{H}_j)$ . Since  $T \in (\rho_1 \oplus \rho_2)(A)'$  we see that  $T_{ij} \in \text{Hom}_A(\rho_i, \rho_j)$ . Thus  $T_{12}, T_{21}$  are 0. We thus see that

$$\begin{bmatrix} \text{Id}_{\mathcal{H}_1} & 0 \\ 0 & 0 \end{bmatrix} \in (\rho_1 \oplus \rho_2)(A)'' = \overline{(\rho_1 \oplus \rho_2)(A)}^{\text{SOT}}$$

where the last equality follows from von Neumann's Double Commutant Theorem. We now prove (iv) by using Kaplansky's Density Theorem. □

We need an analogue of the Lebesgue decomposition.

**Proposition 4.3.** *Let  $A$  be a unital  $*$ -algebra and  $\rho_j: A \rightarrow B(\mathcal{H}_j)$ ,  $j = 1, 2$  be two  $*$ -representations. Then*

$$\rho_1 = \rho_{1,s} \oplus \rho_{1,c}$$

where  $\rho_{1,s} \perp \rho_2$ , and  $\rho_{1,c} \ll \rho_2$ .

*Proof.* By Zorn's Lemma, we can find a maximal family  $(\mathcal{K}_\beta)_{\beta \in B}$  of pairwise orthogonal,  $A$ -invariant, closed, linear subspaces of  $\mathcal{H}_1$  so that  $\rho_1|_{\mathcal{K}_\beta}$  embeds into  $\rho_2$ . Let

$$\mathcal{H}_{1,s} = \mathcal{H}_1 \cap \left[ \bigoplus_{\beta \in B} \mathcal{K}_\beta \right]^\perp.$$

By maximality  $\rho_1|_{\mathcal{H}_{1,s}}$  is singular with respect to  $\rho_2$ . Setting

$$\mathcal{H}_{1,c} = \bigoplus_{\beta \in B} \mathcal{K}_\beta$$

and defining  $\rho_{1,s}, \rho_{1,c}$  by restricting  $\rho_1$  to  $\mathcal{H}_{1,s}, \mathcal{H}_{1,c}$  completes the proof.  $\square$

#### 4.2. Proofs of the Main Applications.

**Theorem 4.4.** *Let  $\Gamma$  be a countable discrete sofic group with sofic approximation  $\Sigma$ . Let  $(X, \mathcal{M}, \mu)$  be a standard probability space with  $\Gamma \curvearrowright (X, \mathcal{M}, \mu)$  by measure-preserving transformations. Let  $\mathcal{H} \subseteq L^2(X, \mu)$  be a closed, linear,  $\Gamma$ -invariant subspace such that  $\mathcal{M}$  is generated by*

$$\{gf^{-1}(A) : g \in \Gamma, f \in \mathcal{H}, A \subseteq \mathbb{C} \text{ is Borel}\}.$$

Suppose that

$$\rho_{\Gamma \curvearrowright (X, \mu)}|_{\mathcal{H}} \perp \lambda_\Gamma.$$

Then

$$h_{\Sigma, \mu}(X, \Gamma) \leq 0.$$

*Proof.* Let  $(f_n)_{n=1}^\infty$  be a  $\|\cdot\|_2$ -dense subset of  $\mathcal{H}$ . Let

$$\Phi: X \rightarrow \mathbb{C}^{\mathbb{N} \times \Gamma}$$

be defined by

$$\Phi(x)(n, g) = f_n(g^{-1}x)$$

and let

$$\nu = \Phi_*(\mu).$$

Let  $\Gamma \curvearrowright \mathbb{C}^{\mathbb{N} \times \Gamma}$  by shifts. Since  $\{gf_n^{-1}(A) : g \in \Gamma, n \in \mathbb{N}, A \subseteq \mathbb{C} \text{ is Borel}\}$  generates  $\mathcal{M}$  we see that  $\Phi$  induces a  $\Gamma$ -invariant isomorphism  $(X, \mu) \cong (\mathbb{C}^{\mathbb{N} \times \Gamma}, \nu)$ . Thus

$$h_{\Sigma, \mu}(X, \Gamma) = h_{\Sigma, \nu}(\mathbb{C}^{\mathbb{N} \times \Gamma}, \Gamma).$$

For  $n \in \mathbb{N}$  let  $Z_n: \mathbb{C}^{\mathbb{N} \times \Gamma} \rightarrow \mathbb{C}$  be defined by

$$Z_n(y) = z(n, e),$$

then  $Z_n \circ \Phi = f_n$ . Thus

$$\rho_{\Gamma \curvearrowright (\mathbb{C}^{\mathbb{N} \times \Gamma}, \nu)}|_{\overline{\text{Span}\{gZ_n : g \in \Gamma, n \in \mathbb{N}\}}} \cong \rho_{\Gamma \curvearrowright (X, \mu)}|_{\mathcal{H}} \perp \lambda_\Gamma.$$

To simplify notation, we will use  $\rho$  for  $\rho_{\Gamma \curvearrowright (\mathbb{C}^{\mathbb{N} \times \Gamma}, \nu)}$ . Let  $\Delta$  be the dynamically generating pseudometric on  $\mathbb{C}^{\mathbb{N} \times \Gamma}$  defined by

$$\Delta(z, w) = \sum_{n=1}^{\infty} 2^{-n} \frac{|z(n, e) - w(n, e)|}{1 + |z(n, e) - w(n, e)|}.$$

As  $\mathbb{C}^{\mathbb{N} \times \Gamma}$  is clearly Polish, we can use  $\mathbb{C}^{\mathbb{N} \times \Gamma}, \Delta, \nu$  to do our computation. Let  $\varepsilon > 0$  be arbitrary, and let  $0 < \eta < \varepsilon$  be arbitrary. Let  $\delta > 0$  be sufficiently small and  $F \subseteq \Gamma$  be a sufficiently large finite set which will depend upon  $\varepsilon, \eta$  in a manner to be determined later. Given  $\phi \in \text{Map}(\Delta, F, \delta, \sigma_i)$  for  $n \in \mathbb{N}$  define  $z_{\phi, n} \in \mathbb{C}^{d_i}$  by

$$z_{\phi, n}(j) = \phi(j)(n)(e)$$

and define  $z_\phi \in (\mathbb{C}^\mathbb{N})^{d_i}$  by

$$z_\phi(j)(n) = z_{\phi,n}(j).$$

Conversely, given  $z \in (\mathbb{C}^\mathbb{N})^{d_i}$  define  $\psi_z: \{1, \dots, d_i\} \rightarrow \mathbb{C}^{\mathbb{N} \times \Gamma}$  by

$$\psi_z(j)(n, g) = z(\sigma_i(g)^{-1}(j))(n)$$

If  $F, \delta$  are chosen carefully, then for all sufficiently large  $i$  we have that

$$\Delta_2(\psi_{z_\phi}, \phi) < \varepsilon.$$

Choose  $N \in \mathbb{N}$  so that

$$2^{-N} < \varepsilon.$$

Since

$$\Gamma \curvearrowright \overline{\text{Span}\{gZ_n : g \in \Gamma, n \in \mathbb{N}\}} \perp \lambda_\Gamma,$$

by Proposition 4.2 we may find a  $\alpha \in \mathbb{C}(\Gamma)$  so that  $\max(\|\lambda(\alpha^* \alpha)\|, \|\rho(\alpha^* \alpha)\|) \leq 1$  and

$$\|\rho(\alpha^* \alpha)Z_n - Z_n\|_2 < \eta, \quad 1 \leq n \leq N$$

$$\|\lambda(\alpha^* \alpha)\delta_e\|_2 < \eta.$$

Let  $M > 0$  be sufficiently large depending upon  $\varepsilon$  in a manner to be determined later. We will assume that  $M$  is large enough so that there exists  $G \in C_c(\mathbb{C})$  with  $\|G\|_\infty \leq M$  and  $G(z) = z$  for  $|z| \leq M$  and

$$\|G \circ Z_n - Z_n\|_2 < \varepsilon \text{ for } 1 \leq n \leq N.$$

Note that  $G$  may be chosen independent of  $\eta$ . As

$$\|\rho(\alpha^* \alpha)\| \leq 1$$

we then have

$$(4) \quad \|\rho(\alpha^* \alpha)G \circ Z_n - G \circ Z_n\|_2 < 3\varepsilon, \text{ for } 1 \leq n \leq N.$$

Let  $L \subseteq C_b(\mathbb{C}^{\mathbb{N} \times \Gamma})$  be sufficiently large in a manner to be determined later. We will assume that

$$L \supseteq \{G \circ Z_1, \dots, G \circ Z_N\}.$$

We will use  $\|\cdot\|_2$  for the  $\ell^2$  norm on  $\{1, \dots, d_i\}$  with respect to the uniform probability measure. By (4) if  $F \subseteq \Gamma, L \subseteq C_b(\mathbb{C}^{\mathbb{N} \times \Gamma})$  are sufficiently large and  $\delta > 0$  is sufficiently small then

$$\|\sigma_i(\alpha)^* \sigma_i(\alpha)G \circ z_{\phi,n} - G \circ z_{\phi,n}\|_2 < 6\varepsilon \text{ for } 1 \leq n \leq N,$$

for all  $\phi \in \text{Map}_\mu(\rho, F, L, \delta, \sigma_i)$  (with the notational conventions introduced after Definition 2.1). Let  $p = \chi_{[1-\sqrt{\varepsilon}, 1+\sqrt{\varepsilon}]}(\sigma_i(\alpha)^* \sigma_i(\alpha))$  (this expression should be interpreted in the sense of functional calculus) then for all  $\phi \in \text{Map}_\mu(\rho, F, L, \delta, \sigma_i)$ ,

$$\begin{aligned} \|pG \circ z_{\phi,n} - G \circ z_{\phi,n}\|_2^2 &= \|\chi_{(\sqrt{\varepsilon}, \infty)}(|\sigma_i(\alpha)^* \sigma_i(\alpha) - 1|)G \circ z_{\phi,n}\|_2^2 \\ &= \langle \chi_{(\sqrt{\varepsilon}, \infty)}(|\sigma_i(\alpha)^* \sigma_i(\alpha) - 1|)G \circ z_{\phi,n}, G \circ z_{\phi,n} \rangle \\ &\leq \frac{1}{\varepsilon} \langle (\sigma_i(\alpha)^* \sigma_i(\alpha) - 1)^2 G \circ z_{\phi,n}, G \circ z_{\phi,n} \rangle \\ &= \frac{1}{\varepsilon} \|(\sigma_i(\alpha)^* \sigma_i(\alpha) - 1)(G \circ z_{\phi,n})\|_2^2 \\ &< 36\varepsilon \end{aligned}$$

and

$$\text{tr}(p) \leq \frac{1}{1-\sqrt{\varepsilon}} \text{tr}(\sigma_i(\alpha)^* \sigma_i(\alpha)).$$

Since

$$\lim_{i \rightarrow \infty} \text{tr}(\sigma_i(\alpha)^* \sigma_i(\alpha)^2) = \tau(\lambda(\alpha)^* \lambda(\alpha)) = \langle \lambda(\alpha)^* \lambda(\alpha) \delta_e, \delta_e \rangle \leq \|\lambda(\alpha)^* \lambda(\alpha) \delta_e\|_2 < \eta$$

we see that for all large  $i$  we have

$$\text{tr}(p) < 2\eta$$

provided  $\varepsilon < \frac{1}{(2015)!}$ . Since  $p$  is an orthogonal projection, we know by Lemma 2.7 we may choose an  $\varepsilon$ -dense subset  $S$  of  $Mp \text{Ball}(\ell^2(d_i, u_{d_i}))$  with

$$|S| \leq \left( \frac{3M + \varepsilon}{\varepsilon} \right)^{4\eta d_i}.$$

If  $L$  and  $M$  are sufficiently large, then

$$C_i = \{1 \leq j \leq d_i : |z_{\phi,n}(j)| \leq M \text{ for } 1 \leq n \leq N\}$$

has

$$|C_i| \geq (1 - \varepsilon)d_i.$$

Define  $w_\phi \in (\mathbb{C}^{\mathbb{N}})^{d_i}$  by

$$w_\phi(j)(n) = G(z_{\phi,n}(j)).$$

For these values of  $F, \delta, L, M$  we have

$$\begin{aligned} \Delta_2(\phi, \psi_{w_\phi})^2 &\leq \varepsilon + \sum_{j=1}^{d_i} \sum_{n=1}^N 2^{-n} \frac{|z_\phi(n, j) - w_\phi(n, j)|}{1 + |z_\phi(n, j) - w_\phi(n, j)|} \\ &\leq 2\varepsilon \end{aligned}$$

as  $w_\phi(n, j) = z_\phi(n, j)$  for  $j \in \{1, \dots, d_i\} \setminus C_i$ . For  $1 \leq n \leq N$  choose  $\xi_{\phi,n} \in S$  so that

$$\|pG \circ z_{\phi,n} - \xi_{\phi,n}\|_2 < \varepsilon$$

and define  $\xi_\phi \in (\mathbb{C}^{\mathbb{N}})^{d_i}$  by

$$\xi_\phi(j)(n) = \chi_{\{1, \dots, N\}}(n) \xi_{\phi,n}(j).$$

Then

$$\begin{aligned} \Delta_2(\psi_{w_\phi}, \psi_{\xi_\phi})^2 &\leq \varepsilon + \sum_{n=1}^N \frac{2^{-n}}{d_i} \sum_{j=1}^{d_i} |G(z_{\phi,n}(j)) - \xi_{\phi,n}(j)| \\ &\leq \varepsilon + \max_{1 \leq n \leq N} \|G \circ z_{\phi,n} - \xi_{\phi,n}\|_2 \\ &\leq 2\varepsilon + 6\sqrt{\varepsilon} \end{aligned}$$

the second line following from the Cauchy-Schwartz inequality.

Thus

$$\Delta_2(\phi, \psi_{\xi_\phi}) \leq \sqrt{2}\sqrt{\varepsilon} + (2\varepsilon + 6\sqrt{\varepsilon})^{1/2},$$

Hence for all large  $i$ ,

$$S_{2\sqrt{2}\sqrt{\varepsilon} + 2(2\varepsilon + 6\sqrt{\varepsilon})^{1/2}}(\text{Map}_\mu(\Delta, F, \delta, L, \sigma_i), \Delta_2) \leq |S|^N \leq \left( \frac{3M + \varepsilon}{\varepsilon} \right)^{4N\eta d_i}.$$

Thus

$$h_{\Sigma, \mu}(\Delta, 4\sqrt{2}\sqrt{\varepsilon} + 4(\varepsilon + 6\sqrt{\varepsilon})^{1/2}) \leq 4N\eta \log \left( \frac{3M + \varepsilon}{\varepsilon} \right).$$

Note that  $\eta$  can be any number in  $(0, \varepsilon)$ . Thus we can let  $\eta \rightarrow 0$  to find that

$$h_{\Sigma, \mu}(\Delta, 4\sqrt{2}\sqrt{\varepsilon} + 4(\varepsilon + 6\sqrt{\varepsilon})^{1/2}) \leq 0$$

since  $\varepsilon > 0$  is arbitrary we find that

$$h_{\Sigma, \mu}(X, \Gamma) \leq 0.$$

□

The space  $\mathcal{H}$  can be much smaller than  $L^2(X, \mu)$ . For example consider the case that  $(X, \mu) = (B, \eta)^\Gamma$  and the action is Bernoulli. For  $f \in L^2(B, \eta)$  and  $g \in \Gamma$ , let  $f_g \in L^2(X, \mu)$  be defined by

$$f_g(x) = f(x(g)).$$

Then we can take

$$\mathcal{H} = \overline{\text{Span}\{f_g : g \in \Gamma, f \in L^2(B, \eta)\}}$$

and one can show that

$$\Gamma \curvearrowright L^2(X, \mu) \ominus \mathcal{H} \cong \ell^2(\Gamma)^{\oplus \infty}.$$

So indeed the space  $\mathcal{H}$  is much smaller than  $L^2(X, \mu)$ .

For the next application, recall that the weak topology on  $\text{Aut}(X, \mu)$  is defined by saying that a basic neighborhood of  $\alpha$  is given by  $U_{A_1, \dots, A_n, \varepsilon}(\alpha)$  where  $A_1, \dots, A_n$  are measurable subsets of  $X$  and  $\varepsilon > 0$  and

$$U_{A_1, \dots, A_n, \varepsilon}(\alpha) = \bigcap_{j=1}^n \{\beta \in \text{Aut}(X, \mu) : |\mu(\beta^{-1}(A_j) \Delta \alpha^{-1}(A_j))| < \varepsilon\}.$$

An action  $\Gamma \curvearrowright (X, \mu)$  is *compact* if there is a compact subgroup  $K \subseteq \text{Aut}(X, \mu)$  (for the weak topology) and a homomorphism  $\pi: \Gamma \rightarrow K$  so that

$$gx = \pi(g)(x)$$

for almost every  $x \in X$ .

**Corollary 4.5.** *Let  $\Gamma$  be a countably infinite discrete sofic group with sofic approximation  $\Sigma$ . Suppose that  $\Gamma \curvearrowright (X, \mu)$  is a compact action. Then  $h_{\Sigma, \mu}(X, \Gamma) \leq 0$ .*

*Proof.* Recall that a unitary representation  $\rho: \Gamma \rightarrow U(\mathcal{H})$  is called *weakly mixing* if

$$0 \in \overline{\rho(\Gamma)}^{WOT},$$

and is *compact* if

$$\overline{\rho(\Gamma)}^{WOT} \subseteq U(\mathcal{H}).$$

It is clear that a compact representation has no nontrivial weakly mixing subrepresentations. It is also well-known that the left regular representation is weakly mixing. Thus if  $\Gamma \curvearrowright (X, \mu)$  is compact, then

$$\rho_{\Gamma \curvearrowright (X, \mu)}^0 \perp \lambda_\Gamma$$

and we may apply Theorem 4.4. □

In particular, note that if  $K$  is a compact group and  $\phi: \Gamma \rightarrow K$  is a homomorphism, then the action  $\alpha$  of  $\Gamma$  on  $K$  given by

$$\alpha(g)(x) = \phi(g) \cdot x$$

has entropy at most zero with respect to any sofic approximation.

**Definition 4.6.** Let  $\Gamma$  be a countable discrete sofic group with sofic approximation  $\Sigma$ . We say that a probability measure-preserving action  $\Gamma \curvearrowright (X, \mu)$  has *completely positive entropy* (with respect to  $\Sigma$ ) if whenever  $\Gamma \curvearrowright (Y, \nu)$  is a factor of  $\Gamma \curvearrowright (X, \mu)$  and  $Y$  is not a one-atom space then  $h_{\Sigma, \nu}(Y, \Gamma) > 0$ .

**Corollary 4.7.** *Let  $\Gamma$  be a countable discrete sofic group with sofic approximation  $\Sigma$ . Suppose that  $\Gamma \curvearrowright (X, \mu)$  is a probability measure-preserving action which has completely positive entropy with respect to  $\Sigma$ . Then  $\rho_{\Gamma \curvearrowright (X, \mu)}^0 \ll \lambda_\Gamma$ .*

*Proof.* By Proposition 4.3, we can write

$$L^2(X, \mu) = \mathcal{H}_1 \oplus \mathcal{H}_2$$

where  $\rho_{\Gamma \curvearrowright (X, \mu)}|_{\mathcal{H}_1} \perp \lambda_\Gamma$  and

$$\rho_{\Gamma \curvearrowright (X, \mu)}|_{\mathcal{H}_2} \ll \lambda_\Gamma.$$

Suppose that  $f \in \mathcal{H}_1$ . Define

$$\Phi: X \rightarrow \mathbb{C}^\Gamma$$

by

$$\Phi(x)(g) = f(g^{-1}x)$$

and let  $\nu = \Phi_*\mu$ , and let  $\Gamma \curvearrowright \mathbb{C}^\Gamma$  be the Bernoulli action. Then  $\Gamma \curvearrowright (\mathbb{C}^\Gamma, \nu)$  is a factor of  $\Gamma \curvearrowright (X, \mu)$ . Set

$$\mathcal{K} = \overline{\text{Span}\{g(f \circ \Phi) : g \in \Gamma\}}$$

then

$$\{g\xi^{-1}(A) : \xi \in \mathcal{K}, \text{Borel } A \subseteq \mathbb{C}, g \in \Gamma\}$$

generates the Borel subsets of  $\mathbb{C}^\Gamma$  up to  $\nu$ -measure zero. Tautologically,

$$\mathcal{K} \cong \overline{\text{Span}\{gf : g \in \Gamma\}} \perp \lambda_\Gamma.$$



Hence by Theorem 4.4 we know that

$$h_{\Sigma, \nu}(\mathbb{C}^\Gamma, \Gamma) \leq 0.$$

Since  $\Gamma \curvearrowright (X, \mu)$  has completely positive entropy this implies that  $(\mathbb{C}^\Gamma, \Gamma)$  is a one-atom space. But this is only possible if  $f$  is constant. Thus  $\mathcal{H}_1 = \mathbb{C}1$  and

$$\rho_{\Gamma \curvearrowright (X, \mu)}^0 = \rho|_{\mathcal{H}_2} \ll \lambda_\Gamma.$$

□

Again the above corollary illustrates the utility in not assuming that  $\mathcal{H} = L^2(X, \mu)$  in Theorem 4.4 but instead just assuming that  $\mathcal{H}$  generates  $(X, \mu)$ . If  $(\mathbb{C}^\Gamma, \nu)$  is as in the above proof we do not a priori know that

$$\Gamma \curvearrowright L^2(\mathbb{C}^\Gamma, \nu) \ominus \mathbb{C}1 \perp \lambda_\Gamma.$$

## REFERENCES

- [1] A. Alpeev. On Pinsker factors for Rokhlin entropy. In *Representation theory, dynamical systems, combinatorial methods. Part XXIV*, volume 432 of *Zap. Nauchn. Sem. POMI*, pages 30–35, St. Petersburg, 2015. POMI.
- [2] J. Bourgain and A. Gamburd. On the spectral gap for finitely-generated subgroups of  $SU(2)$ . *Invent. Math.*, 171(1):83–121, 2008.
- [3] J. Bourgain and A. Gamburd. Uniform expansion bounds for cayley graphs of  $SL_2(\mathbb{F}_p)$ . *Ann. of Math.*, 167(2):625–642, 2008.
- [4] J. Bourgain and A. Yehudayof. Expansion in  $SL_2(\mathbb{R})$  and monotone expanders. *Geom. Funct. Anal.*, 23(1):1–41, 2013.
- [5] L. Bowen. Measure conjugacy invariants for actions of countable sofic groups. *J. Amer. Math. Soc.*, 23:217–245, 2010.
- [6] L. Bowen. A new measure conjugacy invariant for actions of free groups. *Ann. of Math.*, 171(2):1387–1400, 2010.
- [7] L. Bowen. Entropy theory for sofic groupoids I: the foundations. *J. Anal. Math.*, 124(1):149–233, 2014.
- [8] R. Bowen. Entropy for group endomorphisms. *Trans. Amer. Math. Soc.*, 153(401-414), 1971.
- [9] N. Brown and N. Ozawa. *C\*-Algebras and Finite-Dimensional Approximations*. Cambridge University Press, 1994.
- [10] P. Burton. Naive entropy of dynamical systems. *arXiv:1503.06360*.
- [11] L. Ciobanu, D. Holt, and S. Rees. Sofic groups: graph products and graphs of groups. *Pacific Journal of Mathematics*, 271(1):53–64, November 2014.
- [12] A. Connes, J. Feldman, and B. Weiss. An amenable equivalence relations is generated by a single transformation. *Ergodic Theory Dynam. Systems*, 10(4):431–450, 1983.
- [13] J. Conway. *A Course in Functional Analysis*. Graduate Texts in Mathematics. Springer, New York, NY, second edition, 2000.
- [14] A. Dooley and V. Golodets. The spectrum of completely positive entropy actions of countable amenable groups. *J. Funct. Anal.*, 196(1):1–18, 2002.
- [15] K. Dykema, D. Kerr, and M. Pichot. Orbit equivalence and sofic approximation. *arXiv:1102.2556*.
- [16] K. Dykema, D. Kerr, and M. Pichot. Sofic dimension for discrete measurable groupoids. *Trans. Amer. Math. Soc.*, 366(2):707–748, 2013.
- [17] G. Elek and E. Szabo. On sofic groups. *J. Group Theory*, 9(2):161–171, 2006.
- [18] B. Hayes. Fuglede-Kadison determinants and sofic entropy. *arXiv:1402.1135*.
- [19] B. Hayes. Mixing and spectral gap relative to pinkser factors for sofic groups. *arXiv:1509.07839*
- [20] B. Hayes. An  $l^p$ -version of von neumann dimension for Banach space representations of sofic groups. *J. Funct. Anal.*, 266(2):989–1040, 2014.
- [21] A. Ioana. Orbit equivalence and borel reducibility rigidity for profinite actions with spectral gap. *arXiv:1309.3026*.
- [22] D. Kerr. Sofic measure entropy via finite partitions. *Groups Geom. Dyn.*, 7(617-632), 2013.
- [23] D. Kerr and H. Li. Soficity, amenability, and dynamical entropy. *Amer. J. Math.*, 135(3):721–761, 2013.
- [24] D. Kerr and H. Li. Bernoulli actions and infinite entropy. *Groups Geom. Dyn.*, 5:663–672, 2011.
- [25] D. Kerr and H. Li. Topological entropy and the variational principle for actions of sofic groups. *Invent. Math.*, 186:501–558, 2011.
- [26] D. Kerr and H. Li. Combinatorial independence and sofic entropy. *Comm. Math. Stat.*, 1(2):213–257, 2014.
- [27] J. Kieffer. A generalized Shannon-McMillan theorem for the action of an amenable group on a probability space. *Ann. Prob.*, 3(6):1031–1037, 1975.
- [28] L. Bowen and H. Li. Harmonic models and spanning forests of residually finite groups. *J. Funct. Anal.*, 263(7):1769–1808, 2012.
- [29] H. Li. Compact group automorphisms, addition formulas and Fuglede-Kadison determinants. *Ann. of Math.*, 176(1):303–347, 2012.
- [30] H. Li. Sofic mean dimension. *Adv. Math.*, 244:570–604, 2014.
- [31] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.
- [32] G. Margulis. Some remarks on invariant means. *Monatsh. Math.*, 90(3):233–234, 1980.
- [33] T. Meyerovitch. Positive sofic entropy implies finite stabilizer. *arXiv:1504.08137*.
- [34] D. Ornstein and B. Weiss. Ergodic theory of amenable groups I. The Rokhlin lemma. *Bull. Amer. Math. Soc.(N.S.)*, 2(1):161–165, 1980.

- [35] D. Ornstein and B. Weiss. Entropy and isomorphism theorems for actions of amenable groups. *J. Anal. Math.*, 48:1–141, 1987.
- [36] L. Paunescu. On sofic actions and equivalence relations. *J. Funct. Anal.*, 261(9):2461–2485, November 2011.
- [37] S. Popa. On the superrigidity of malleable actions with spectral gap. *J. Amer. Math. Soc.*, 21(4):981–1000, 2008.
- [38] S. Popa. Independence properties in subalgebras of ultraproduct  $\text{II}_1$  factors. *J. Funct. Anal.*, 266(9):5818–5846, 2014.
- [39] A. Selberg. On the estimation of Fourier coefficients of modular forms. In *Proc. Sympos. Pure Math*, volume 8, pages 1–15, Providence, RI, 1965. American Mathematical Society.
- [40] B. Seward. Finite entropy actions of free groups, rigidity of stabilizers, and a Howe–Moore type phenomenon. *To appear in J. Anal. Math.*
- [41] B. Seward. Krieger’s finite generator theorem for ergodic actions of countable groups I. *arXiv:1405.3604*.
- [42] J. Sinaĭ. On a weak isomorphism of transformations with invariant measure. *Mat. Sb. (N.S.)*, 63(105):23–42, 1964.
- [43] D. Sullivan. For  $n > 3$  there is only one finitely additive rotationally invariant measure on the  $n$ -sphere on all Lebesgue measurable sets. *Bull. of the AMS*, 4(1):121–123, 1981.
- [44] V. Varadarajan. Measure on topological spaces. *Mat. Sb. (N.S.)*, 55(97):35–100, 1961.

STEVENSON CENTER, NASHVILLE, TN 37240

*E-mail address:* `benjamin.r.hayes@vanderbilt.edu`